Using Association Rule to Analyze Hypermarket Customer Purchase Patterns

Ivan Diryana Sudirman, Ronny Samsul Bahri, Iston Dwija Utama, Chyntia Ika Ratnapuri
Entrepreneurship Department, BINUS Business School Undergraduate Program
Bina Nusantara University
Bandung, West Java, 40181, Indonesia
ivan.sudirman@binus.edu, ronny.bahri@binus.ac.id, iston.utama@binus.edu,
chyntia.ratnapuri@binus.ac.id

Abstract

Understanding consumer buying behavior is compulsory in business. By understanding the purchasing behavior, a supermarket can effectively make decisions regarding many things, such as the placement of goods or carry out appropriate promotional activities. One of affected the consumer purchasing pattern is budget, the amount of the budget is undoubtedly related to the income of the consumer. This study aims to reveal changes in consumer purchasing patterns in three periods in one month, namely the first ten days in a month, the second 10 days, and the last ten days in a month. Thus, to answer these objectives, data mining is used with the Association Rules technique. The CRISP-DM method is used as a guide in the steps towards the data mining process. The results indicate that consumer purchase patterns are different in all three periods in one month. In the first ten days of the month, the association between instant noodles and chicken eggs is quite complicated. In the second 10 days, a new association appeared, namely cooking oil and broiler chicken, and on the third 10 days, only one association rule appeared. This study also provides a snapshot of another possibility related to the use of association rules.

Keywords
Market basket analysis, FP Growth, Transaction data, Association rule

1. Introduction
In recent times, the high quality of information available through internet technology has opened up the opportunity to develop and implement specific parts of management information systems, including fast development of methodologies such as data mining, data warehouse, and business analytics. Data mining involved in implementing a wide range of different methods of statistical and engineering to recognize unseen interaction from vast amounts of data. Using data mining techniques, it creates models that can assist in decision making. Data mining can reveal the patterns of a dataset that can be used to understand consumer behavior. Then used as decision-making as Kotler and Armstrong (2010) stressed that marketing information might be used to generate understanding from clients to make more robust marketing choices.

Data mining can be combined with other applications such as CRM to produce benefits for the company. Rygielski et al. (2002) considered Customer Relationship Management (CRM) and Data Mining combination as an essential resource for securing substantial competitive benefits in recognizing profitable consumers, predicting future behavior as well as encouraging businesses to take strategic, knowledge-based decisions.

Decision making and recognizing consumer behavior are important and difficult problems for companies to retain their competitiveness in the market. Technological advancement also enables businesses to address the customer needs better and wants. Data mining methods, perhaps the most promising method for analyzing large amounts of data, take to facilitate effectiveness in decision making. Data mining technology can be used effectively to identify trends and to automate the complex behavior of customer purchases for the purchasing of particular goods (Vahidi Roodpishi & Aghajan Nashtaei, 2015).

Recognizing consumer behavior is essential because, in this situation, marketers must understand what motivates consumer purchasing decisions. By understanding how individuals decide on a product, marketers could perhaps

enclose the gaps in the market and recognize the goods that are needed as well as the outdated goods. To recognize the behavior of customers, marketers must create a superior customer experience to stimulate and attract buyers. Much research has shown that creating customer experience can influence customer decision-making and can be a competitive advantage for the company (Schmitt, 2010). In the retail business, understanding of consumer purchasing patterns is fundamental. Understanding the behavior patterns of consumers can be the basis in various decision making such as the placement of goods, services, promotional activities. Good retailers must have a package of personalized services that satisfies and is appropriate for consumers to be active and succeed in gaining a competitive advantage (Ha et al., 2002).

Many factors affect consumer purchasing patterns, one of which, and perhaps the most influential is the budget owned by these consumers. Before making a purchase decision, customers will take their budget into account. Buyers go shopping as feasible and anticipate service, perhaps in the strangest times. It is the responsibility of the company to satisfy such needs by recognizing a purchase pattern and matching its service with the purchase time and frequency. (Radu,2019). To help marketers understand the consumer purchasing patterns behavior, we can use sensory marketing that consists of five senses which play roles to build the customer experience such as sound, taste, scent, touch, and sight (Hultén et al., 2009).

The budget that is owned by consumers is very likely influenced by income. Income is usually obtained at the beginning of the month and slowly decreases at the end of the month. So there is a possibility that consumers' shopping patterns can change along with the current day of the month. Based on this, we can divide in one month into three parts, namely the first ten days, the second 10 days, and the last ten days in a month, to see if there is a change in shopping patterns from consumers.

1.1 Objectives

The purpose of this study is to reveal the shopping patterns of consumers in retail business in three terms in one month. It is hoped that this research will contribute to decision making related to a similar industry. For this purpose, market basket analysis using association rules is one of the most popular methods.

2. Literature Review

In this study, a market basket analysis using the association rules will be conducted. Because in the retail market, most of the sales are made on compulsion. Market basket analysis provides insight as to what a consumer would have purchased if they had the idea. Market Basket Analysis (MBA) can indeed be used as an initial phase in determining the location and advertising of products within a retailer (Ünvan, 2020). Market Basket Analysis (MBA) is a collection of statistical association formulas that enables marketers to have better understand and, necessarily, meet the needs of customers through a focus on retail buying behavior. The MBA indicates what types of items most often appear together within transactions.

The association rule approach is often used to perform a market basket analysis. Association rules are a data mining tool for exploring behaviors of frequent itemsets, such as goods in a retailer that are frequently purchased at about the same moment by a consumer. Association rules are chosen based on its ability to identify hidden patterns among significant numbers of data, primarily from transactional databases (Liao & Chang, 2016).

The basic algorithm used for the association rules is the apriori algorithm developed by Agrawal and Srikant (1994). This algorithm works by scanning dataset, examines every possible rule, and afterward maintains only such rules which have Support and Confidence higher than the predetermined lowest value as the interesting ones. The apriori algorithm for constructing association rules for transactional data comprises of the two following steps. First is discovering all frequent itemsets in a transactional database that meet the requirements support, and second is generating association rules for frequent items that meet the minimum Confidence.

Another algorithm for association rules is FP-growth. It is a well-known popular mining algorithm. It scans the database just twice and discovers all standard frequent itemsets efficiently, especially in comparison to the Apriori algorithm. FP-growth has three strengths. First, FP-growth condenses the entire database into a relatively small data structure (FP-tree), leading with only two times scanning the database. Second, it builds up a frequent pattern-growth formula to avoid generating enormous candidate itemsets. Third, it produces the detail layers tree to explore

frequent itemsets, reducing computational complexity. Based on the experimental results, FP-growth is quicker than the Apriori algorithm and several methods of frequent mining items (Lin et al., 2011).

According to Chauhan (2019), Support is for the standard popularity of a product. The Support for item A in scientific terms is the ratio of transactions containing A to the total transaction number.

$$Support\ (A) = \frac{(Transactions\ involving\ A)}{(Total\ transaction)}$$

The closer of Support value to one, the better. It is mean that the item is frequently appear in the transaction. As Confidence, which is the likelihood that consumers bought both item A and item B. The total of transactions involving item A and item B is distributed by the number of transactions involving item B.

$$Confidence\ (A \Rightarrow B) = \frac{(Transactions\ involving\ both\ A\ and\ B)}{(Transactions\ involving\ only\ A)}$$

So, the likelihood a consumer purchase both item A and item B together is the Lift value times higher than a possibility when buying alone.

- Lift (A => B) = 1 means that there is no correlation within the itemset.
- Lift (A => B) > 1 means that there is a positive correlation within the itemset, i.e., products in the itemset, A, and B, are more likely to be bought together.
- Lift (A => B) < 1 means that there is a negative correlation within the itemset, i.e., products in itemset, A, and B, are unlikely to be bought together.

Association rule-based algorithms are described as a two-step approach; the first step is Frequent Itemset creation, which includes all frequent Items as given in the minimum Support registered. The second step is creating the rule according to minimum of confidence level.

From the literature that can be studied, generally, the data used in Association Rule is transaction data. Like the research conducted by Setiabudi et al. (2011), the market basket analysis methodology has been applied in the X minimarkets. A search for the Apriori algorithm frequents items that regularly appear in the archive and pairs of items in a single transaction. Frequently chosen items should include a pair of items that surpass the minimum support. Frequent products, after processing, exceeding the required support yield association rules. The findings of the tests indicate that the framework will collect details about what kinds of products consumers often buy at the same time, in compliance with the Association rules about hybrid measurements.

Susac and Has (2015) tries to evaluate objective measures like Support, Confidence, and Lift with a subjective method based on the assortment of human intelligence to recover exciting rules from a dataset acquired from a big Croatian retail store. The categorized rules of association were used to enhance extraction rule efficiency. Ozc¸alıcı (2017) Genetic algorithm is utilized to select the critical components. 252645 advertisements are evaluated for this objective. There are 139 added features for each advertisement promotion. Different models are analyzed using 5, 10, 15, and 20 components of genetic algorithms. The highest execution projection for the off-sample experimentation is 65.67 percent. Kurniawan (2018) estimate transaction data purchasers shopping basket on the transaction data of Business Centre (BC) UIN Malang store. The result shows that the Confidence value median of 46.69% from the Support percentage of 1.78% and the generated rules are 30 rules.

3. Methods

To increase the rate of success in developing data mining tasks, scientists and experts have put in place a range of methodologies (workflows or simple, stage-by-stage procedures). The Cross-Industry Standard Data Mining Process (CRISP-DM) was developed by a European group of companies in the early nineties as an appropriate data mining process (Sharda et al., 2018). Even though members in the development of CRISP-DM were interested in particular hardware and software, the methodology was proposed independently of any particular device. It was designed in such a way as to be conceptual, one that could be used independently of almost any particular tool or type of data (North, 2012). The process consists of six steps or phases:

Phase 1. Business Understanding. A reliable business summary of the research that will be conducted will assist in addressing the situation in responding to management's need for new information. In this study, the business that was observed was retail that had long been operating in Bandung. One of the biggest supermarkets in Bandung, with quite several branches in the city. This supermarket is one of the local supermarkets that must compete with foreign supermarkets such as Carrefour. Competitive advantage is significant to be able to survive amid increasingly fierce retail competition.

Phase 2. Data Understanding. The primary purpose of data mining is to understand data from multiple reputable sources in the sense of corporation recognition. Numerous vital points must be considered throughout the identifying and collection step of the data. To fully understand the most necessary details, the researcher must be clear about the context of the dataset. The data used in this study is transaction data in January. More than 32,000 transaction data consisting of 24 attributes are available for study. The 24 attributes of existing data consist of transaction status, transaction date, time, store, PLU, PLU Description, Quantity, Gross, Net, Busdate, Post Number, Transaction number, article code, division code, category code, category description, sub-category, class code, class description, subclass, subclass description. Of these attributes that will be used in this study is PLU Description, which is a description of the items purchased, based on the transaction number. The contraction number will be divided into three terms in one month, namely the first ten days, the second 10 days, and the last ten days.

Phase 3. Data Preparation. The data preparation aims to examine the data available for mining and analysis, as indicated in the previous section. Available data is organized by items purchased. For example, on January 1, an item was sold as many as four items and recorded as transaction number 35131. So that in several lines can have the same transaction number but have different items. This situation happens because, in the same transaction number, there are many different purchases of goods. The structure of presenting raw data like this becomes a challenge in the data preparation step. Transaction number attribute and PLU description are taken then pivot in such a way as to produce a new table with items purchased as attributes as shown in Table 1.

Table 1. Example of The Dataset

TRANS_DA	TRANS_NO	BABY PA	PEPSODEN	BELIMBIN	LUWAK SC	GILLETTE C	SO KLIN SC
1/1/2017	20800	0	0	0	0	0	0
	20802	0	0	0	0	0	0
	20803	0	0	0	0	0	0
	20804	0	0	0	0	0	0
	20805	0	0	0	0	0	1
	20806	0	0	0	0	0	0
	20807	0	0	0	0	0	0
	20809	0	0	0	0	0	0
	20810	0	0	0	0	0	0

As presented in Table 1, data is arranged based on transaction number, the products purchased by consumers become attributes. At first, the number of attributes is vast; given the limited computing resources, only products purchased more than 100 products in a month will be used. So that the total number of data is 10.608, and the total number of attributes to be processed is 93 attributes, ranging from Baby Pakcoy to local egg.

Phase 4. Model Building. The various models generated during the model creation process are also evaluated and examined. Despite a thorough understanding of the technical application of data mining, an exploration and evaluation approach should be implemented to obtain the "correct" approach in a specific way for some specific purpose. In this study, the method used is the association rule. The data processing uses RapidMiner with the FP Growth operator. RapidMiner is a data mining program that is quite widely used because users do not have to be able to code. RapidMiner uses operator boxes that are connected in such a way as to process data. According to RapidMiner documentation, The FP-Growth algorithm is an efficient algorithm for frequently calculating co-occurring items in a transaction database. FP-tree data structure can be efficiently created, compressing the data so much that, in many cases, even large databases will fit into the main memory. As explained previously, association rules are then formed and studied.

Phase 5. Testing and Evaluation. The accepted model is verified and reviewed for validity and consistency. This stage tests how well and to what extent the proposed model or template suits the objectives and goals. An alternative is to test a model that is part of a real-world system if funds and budgets are available. The adaptation of this step to the association rule is to look at different measurement results. Generally, what is seen is Support and Confidence, but other measurements will also be considered in this study, such as Lifts and other measurements.

The sixth step is deployment will not be discussed in this paper because this step is mostly done by the client, not by the researcher.

4. Results and Discussion

The association rules method in this study uses FP Growth. After the data is imported into RapidMiner, then it is called using the Retrieve operator. Furthermore, the Select Attributes operator is used to select which attributes will be processed, as shown in Figure 1.

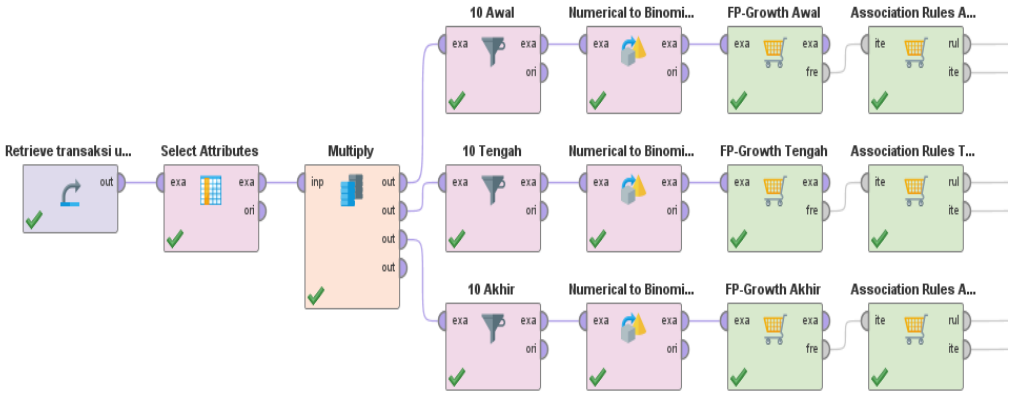

Figure 1. Process in RapidMiner

The Multiply operator is used to distribute data to three paths, as shown in Figure 1. Next, the three example range filter operators are used to divide data by date into the first ten days, the second 10 days, and the last ten days. So that data can be processed using FP Growth, the data must first be changed to Binominal. Therefore, the Numerical to Binominal operator is used so that the form of data is no longer the numbers 1 and 0 but becomes True and False. Then each path uses the FP Growth operator with the same parameters. At first, the minimum Support was used as the lower limit, but after it was processed, association rules could not appear even though the minimum Confidence had been reduced by 50%. Therefore, the minimum support parameter in FP Growth is changed to the frequency with a minimum frequency of 100. In the operator association rules, the parameter used in each operator is minimum Confidence of 50%. After all, settings are ready; the process in Figure 1 is then run. With a simple method like that, we get three results for each term in one month as desired.

Table 2. Frequent Itemsets First 10 Days

Size	Support ↓	Item 1	Item 2
1	0.192	TELUR ,A NEGERI TBG	
1	0.106	INDOMIE AYAM BAWANG	
1	0.097	INDOMIE GR SPC SAUS	
1	0.077	FORTUNE POUCH 2LT	
1	0.073	ANGGUR MERAH	
1	0.066	YAKULT	
1	0.064	DRGN RED DRAGONFRUIT	
1	0.058	CEMARA POUCH 2L	
2	0.053	INDOMIE AYAM BAWANG	INDOMIE GR SPC SAUS
1	0.051	PH GULA LOKAL 1KG	
1	0.050	SUNLIGHTPPLIMEPCH800	
1	0.046	PUCUKHARUMJASMINE350	
1	0.044	FRISIAN.F UHTCHOC900	
1	0.042	SANIA P,OUCH 2 LT	
1	0.042	TESSA TP02 SP260	
1	0.041	FRISIAN.F UHTF.CR900	
2	0.041	TELUR ,A NEGERI TBG	INDOMIE GR SPC SAUS
2	0.040	TELUR ,A NEGERI TBG	INDOMIE AYAM BAWANG

In the first ten days, the frequency can be seen in Table 2, with Support is sorted from biggest to lowest. The table shown is only part of the whole table, but we can see that the highest Support is 19.2% for domestic egg items. The first product pair to appear in Table 2 is Indomie GR SPC Sauce with Indomie Ayam Bawang. Support value for this product pair is 0.053.

Table 3. Frequent Itemsets 2nd 10 Days

Size	Support ↓	Item 1	Item 2
1	0.208	TELUR ,A NEGERI TBG	
1	0.097	INDOMIE AYAM BAWANG	
1	0.083	INDOMIE GR SPC SAUS	
1	0.072	SANIA P,OUCH 2 LT	
1	0.068	ANGGUR MERAH	
1	0.051	YAKULT	
1	0.051	JERUK SHANTANG	
1	0.047	DRGN RED DRAGONFRUIT	
2	0.045	INDOMIE AYAM BAWANG	INDOMIE GR SPC SAUS
2	0.040	TELUR ,A NEGERI TBG	INDOMIE AYAM BAWANG
1	0.038	APEL FUJI RRT	
1	0.038	FORTUNE POUCH 2LT	
2	0.037	TELUR ,A NEGERI TBG	INDOMIE GR SPC SAUS
1	0.033	LENGKENG BANGKOK	
1	0.033	PH GULA LOKAL 1KG	
1	0.032	BREX RJK CIRENG 20'S	

In the second 10 days, the frequency can be seen in Table 3. The three items with top Support are still the same, and the difference starts to be seen in the next item. While item 2 that first appeared was the same, namely Indomie GR SPC Saus (instant noodle) with item 1 being Indomie Ayam Bawang (instant noodle).

Table 4. Frequent Itemsets Last 10 Days

Size	Support ↓	Item 1	Item 2
1	0.210	TELUR ,A NEGERI TBG	
1	0.132	INDOMIE GR SPC SAUS	
1	0.102	INDOMIE AYAM BAWANG	
1	0.071	ANGGUR MERAH	
1	0.071	JERUK SHANTANG	
1	0.066	CEMARA POUCH 2L	
1	0.054	YAKULT	
1	0.053	DRGN RED DRAGONFR...	
1	0.048	SUNLIGHTPPLIMEPCH8...	
2	0.048	INDOMIE GR SPC SAUS	INDOMIE AYAM BAWANG
1	0.047	MAMA PP LIME REF 800	
2	0.046	TELUR ,A NEGERI TBG	INDOMIE GR SPC SAUS
1	0.044	LENGKENG BANGKOK	
2	0.042	TELUR ,A NEGERI TBG	INDOMIE AYAM BAWANG

In the last ten days, the frequency can be seen in Table 4. Just as before the three items with top Support are still the same, the difference starts to be seen in the next item. While item 2 that first appeared was Indomie Ayam Bawang (instant noodle) with item 1 being Indomie GR SPC Sauce (instant noodle), which was slightly different from the previous Table 3.

Table 5. Association Rules First 10 Days

Premises	Conclusion	Support	Confidence	Lift	Convictior
INDOMIE AYAMSPPCS	INDOMIE GR SPC SAUS	0.011874	0.55263158	5.698711	2.018527
TELUR ,A NEGERI TBG, INDOMIE AYAM BAWANG	INDOMIE GR SPC SAUS	0.022053	0.54545455	5.624702	1.986655
INDOMIE GR SPC SAUS	INDOMIE AYAM BAWANG	0.052587	0.54227405	5.101126	1.952468
SYR BWG PUTIH CURAH	TELUR ,A NEGERI TBG	0.011026	0.58208955	3.036653	1.934175

After going through the FP Growth operator, the association rules can be formed. Table 5 presents the association rules that were formed in the first ten days. The support score obtained for each association rules is quite low, ranging from 0.011 to 0.053. Support shows how often relationships between items appear (Agrawal et al., 1993). Although Support is often used as an evaluation measurement for association rules, Support has a weakness. This measurement is vulnerable to rare item problem. (Rage & Krishna Reddy, 2009). Another measure that can be seen is the Lift (Brin et al., 1997)., where if the Lift is more significant than one, then every Y item purchased, then item X is also purchased. Unlike Support, Lift is not affected by rare item problems (Hahsler, 2020). However, Lift is susceptible to noise in small databases; fortunately, in this study, the data used is quite large (35131 transaction data). Conviction measures the implication strength of the rule from statistical independence (Brijs et al., 2003). Also, it can be interpreted as the ratio of the expected frequency that the event occurs without the b. Conviction is another measure proposed to tackle some of the weaknesses of Confidence and lift. Unlike Lift, Conviction is sensitive to rule direction. Like Lift, conviction values far from 1 indicate interesting rules (Azevedo & Jorge, 2007). In an association rule framework, Confidence is a standard measure; Confidence ranges from 0 to 1 with the closer to 1, the better.

Therefore, the association rules that were formed in the first ten days are quite good. Support is small, but Confidence is sufficient, while Lift and Conviction show that the association formed is quite interesting. It seems that the weakness of Support is seen in this study, this might be due to the large number of transactions.

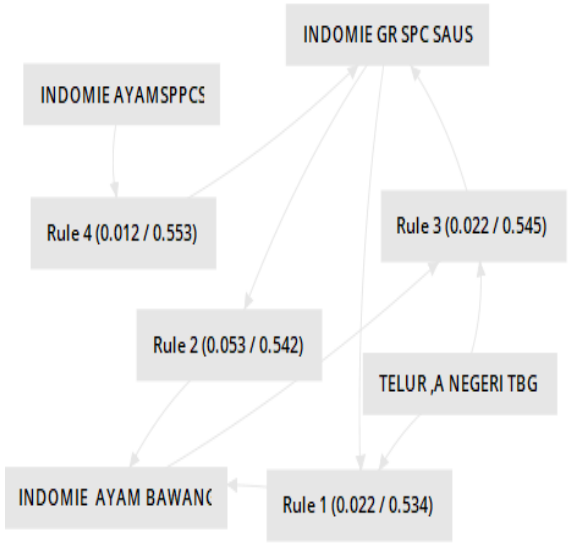

Figure 2. FP Tree First 10 Days

FP Tree in Figure 2 created using the Lift as a minimum criterion and formed four rules. Rule 1 refers to the association of Telur Ayam (Chicken Eggs), Indomie GR SPC Saus (instant noodle) with Indomie Ayam Bawang (instant noodle). Rule 2 is an association between Indomie GR SPC Saus and Indomie Ayam Bawang, and both are instant noodles from the same manufacturer. Rule 3 is Indomie Ayam Bawang, Chicken Egg with Indomie GR SPC Saus. Rule 4 is Indomie Ayam SPPCS with Indomie GR SPC Saus. In general, this shows that many consumers purchase Indomie variant with the favorite variant is Indomie GR SPC, Indomie Ayam Bawang, and Indomie Ayam SPPC, and most of the time, they purchased it along with Chicken Eggs. One of the favorite dishes in Indonesia is instant noodles that are cooked with boiled eggs; uniquely, it was also revealed by this association. Perhaps association rules can also reveal the favorite dish in a society.

In Table 6, for the second 10 days in the dataset, a high Lift score is there in the Filma Pouch 2000 ml with Ayam Broiler Utuh (whole broilers chicken). Overall, the measurements obtained, reveal similar value, including small Support score range between 0.012 to 0.045. While the Confidence could reach 62.3% with Conviction, which is above one and the Lift score, which is also quite far from 1, this shows that there is an interesting association in the data, even though the Support score is low.

Table 6. Association Rules 2nd 10 Days

Premises	Conclusion	Supp...	Confi...	LaPl...	Gain	p-s	Lift	Con...
INDOMIE GR SPC SAUS	INDOMIE AYAM BAWANG	0.045	0.545	0.965	-0.120	0.037	5.628	1.983
TELUR ,A NEGERI TBG, INDOMIE GR SPC ...	INDOMIE AYAM BAWANG	0.021	0.569	0.985	-0.053	0.017	5.884	2.097
FILMA POUCH 2000 ML	AYAM BROILR UTUH/EK	0.012	0.623	0.993	-0.027	0.012	22.711	2.581
KARA SUN SNTN KLP65M	TELUR ,A NEGERI TBG	0.012	0.631	0.993	-0.025	0.008	3.030	2.144

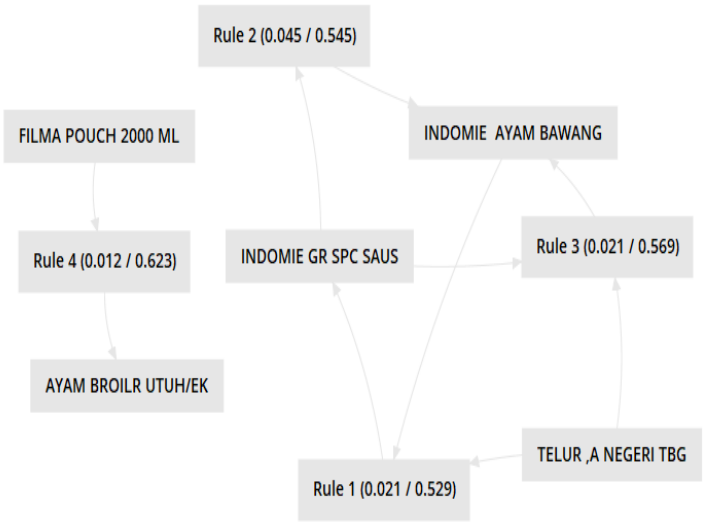

Figure 3. FP Tree 2nd 10 Days

In Figure 3, using the minimum elevator as a criterion, FP Tree with four rules is obtained. Rule 1 is Indomie Ayam Bawang, Telur Ayam with Indomie GR SPC Saus. Rule 2 is Indomie GR SPC Saus with Indomie Ayam Bawang. Rule 3 is Indomie GR SPC Saus, Telur Ayam, with Indomie Ayam Bawang. Rule 4 is a Filma Pouch 2000 ml (cooking oil) with Ayam Broiler Utuh (whole broiler chicken). There is a difference in the establishment of a rule association in the second 10 days, the rule association between instant noodles and chicken eggs persists. However, a new association emerges, namely cooking oil with the whole broiler chicken.

The last ten days from the dataset can be seen in Table 7, where there are only one premise and conclusion relationship. The Support score is also small as before, but the Lift is higher than one and Conviction is also greater than 1. Confidence is 51.4%, but LaPlace is close to 1, this shows that there are interesting associations in the data.

Table 7. Association Rules Last 10 Days

Premises	Conclusion	Sup...	Conf...	LaP...	Gain	p-s	Lift	Con...
BOGASARISEGI3BIRU...	TELUR ,A NEGERI...	0.010	0.514	0.990	-0.030	0.006	2.442	1.624

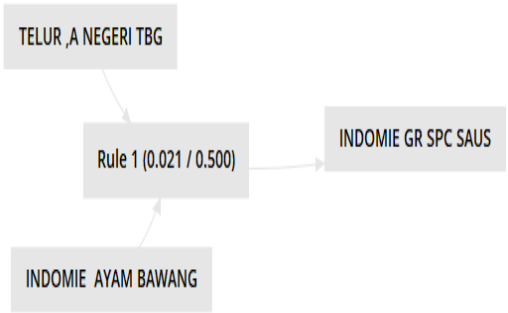

Figure 4. FP Tree Last 10 Days

Figure 4 shows the FP Tree with the minimum criteria is Lift. Only 1 rule was formed in the last ten days, namely Telur Ayam, Indomie Ayam Bawang with Indomie GR SPC Saus. The chicken egg association with instant noodles is still there even though this time is not as complicated as before

5. Conclusion

Companies engaged in retail business must be able to understand their consumer behavior. This study tries to reveal changes in consumer behavior patterns when shopping at supermarkets. The transaction is divided into three terms in one month. The division into 3 periods is done because the decision of consumers to buy products depends on the budget. Transaction data are taken from a local retail company that is which is very famous locally thus the number of goods and the number of transactions is quite large. The dataset is prepared in such a way and then processed using the data mining method, namely Association rules using RapidMiner. FP Growth technique was chosen in this study by considering various advantages such as processing speed. Results show a low Support score at all three time periods in one month. The low value of Support may occur because of the very large number of transactions. As explained earlier that the data used is transaction data from a very well-known hypermarket. Hypermarkets provide a variety of products, local and imported. This causes the number of transactions per day to be very high, thus submerging the number of items purchased simultaneously. However, other association rules measurements such as Lift and Conviction shows that there is an interesting relationship in the data. This association is also supported by a satisfactory Confidence score.

The results of the association rules show that in the three-period tested, and there are changes in the purchase patterns of consumers. In the first ten days, the association between instant noodles and eggs was revealed. So that in the first ten days, the placement of the two items will be better when close together. Promotion between the two products is also possible, or promotion between one variant of instant noodles and other instant noodles that are often bought together.

In the second 10 days, the association between instant noodles and eggs is still present but a bit simpler. In this period, what is interesting is the emergence of the association of cooking oil with broiler chickens with a very high Lift score. Therefore, in the second 10 days of the month, it is possible to place those pair close enough between cooking oil and broiler chicken. Also, it is possible to conduct the promotion of the two items in the form of a bundle, package, or other. In the last ten days of the month, only one association was formed, namely one variant of instant noodles and chicken eggs with another variant of instant noodles.

It seems that the pattern of the association of instant noodles with chicken eggs looks complicated at the beginning of the month and becomes simpler at the end of the month. The difference in purchasing patterns in these three periods from the complex to the simple pattern, shows that indeed the budget is one of the factors in making decisions when buying. The first ten days the consumer still has a flexible budget, while the last ten days the consumer budget is running low.

This study also shows that the company can divide one operational month into three different periods which management can arrange the placement of goods or carry out promotional activities. Besides, another interesting finding is the possibility of association rules in this dataset that can reveal favorite foods in a community. This finding can be used as a basis for further research or decision making in business.

This study only uses data for one month, and at the beginning of the year, it is possible in other months the pattern can change. Also, the attributes used are only products with more than 100 transactions a month. This limitation is due to the limited computational resources owned by researchers. Thus, there may be patterns that are not revealed but turned out to be very interesting to explore. Future research can be done using data in different months, especially in the month of Ramadhan (Fasting), where the consumption patterns of the community will change drastically compared to the regular month.

References

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 207–216. https://doi.org/10.1145/170035.170072

Azevedo, P. J., & Jorge, A. M. (2007). Comparing Rule Measures for Predictive Association Rules. In J. N. Kok, J. Koronacki, R. L. de Mantaras, S. Matwin, D. Mladenič, & A. Skowron (Eds.), *Machine Learning: ECML 2007* (Vol. 4701, pp. 510–517). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74958-5_47

Brijs, T., Vanhoof, K., & Wets, G. (2003). *Defining Interestigness for Association Rules*. http://sci-gems.math.bas.bg:8080/jspui/handle/10525/964

Brin, S., Motwani, R., Ullman, J. D., & Tsur, S. (1997). Dynamic itemset counting and implication rules for market basket data. *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data*, 255–264. https://doi.org/10.1145/253260.253325

Chauhan, N. S. (2019). Market Basket Analysis: A Tutorial. *KDnuggets*. https://www.kdnuggets.com/market-basket-analysis.html/

Ha, S. H., Bae, S. M., & Park, S. C. (2002). Customer's time-variant purchase behavior and corresponding marketing strategies: An online retailer's case. *Computers & Industrial Engineering*, *43*(4), 801–820. https://doi.org/10.1016/S0360-8352(02)00141-9

Hahsler, M. (2020). *A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules*. A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules. https://michael.hahsler.net/research/association_rules/measures.html

Hultén, B., Broweus, N., & Dijk, M. van. (2009). *Sensory marketing*. Palgrave Macmillan.

Kotler, P., & Armstrong, G. M. (2010). *Principles of Marketing*. Prentice Hall.

Kurniawan, F., Umayah, B., Hammad, J., Nugroho, S. M. S., & Hariadi, M. (2018). *Market Basket Analysis to Identify Customer Behaviors by Way of Transaction Data*. 6.

Liao, S., & Chang, H. (2016). A rough set-based association rule approach for a recommendation system for online consumers. *Information Processing & Management*, *52*(6), 1142–1160. https://doi.org/10.1016/j.ipm.2016.05.003

Lin, K.-C., Liao, I.-E., & Chen, Z.-S. (2011). An improved frequent pattern growth method for mining association rules. *Expert Systems with Applications*, *38*(5), 5154–5161. https://doi.org/10.1016/j.eswa.2010.10.047

North, M. (2012). *Data mining for the masses*. Global Text Project.

Özçalıcı, M. (2017). Predicting Second-Hand Car Sales Price Using Decision Trees and Genetic Algorithms. *Alphanumeric Journal*, *5*(1), 103–114.

Radu, Valentin. (2019). Consumer behavior in marketing—Patterns, types, segmentation—Omniconvert Blog. (2019, November 26). *ECOMMERCE GROWTH Blog*. https://www.omniconvert.com/blog/consumer-behavior-in-marketing-patterns-types-segmentation.html

Rage, U., & Krishna Reddy, P. (2009). *An Improved Multiple Minimum Support Based Approach to Mine Rare Association Rules*. 340–347. https://doi.org/10.1109/CIDM.2009.4938669

Rakesh Agrawal & Ramakrishnan Srikant. (1994). Fast Algorithms for Mining Association Rules. *Proceedings of 20th International Conference on Very Large Data Bases VLDB '94*, 13.

Rygielski, C., Wang, J.-C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, *24*(4), 483–502. https://doi.org/10.1016/S0160-791X(02)00038-6

Schmitt, B. (2010). Experience Marketing: Concepts, Frameworks and Consumer Insights. *Foundations and Trends® in Marketing*, *5*(2), 55–112. https://doi.org/10.1561/1700000027

Setiabudi, D. H., Budhi, G. S., Purnama, I. W. J., & Noertjahyana, A. (2011). Data mining market basket analysis' using hybrid-dimension association rules, case study in Minimarket X. *2011 International Conference on Uncertainty Reasoning and Knowledge Engineering*, *1*, 196–199. https://doi.org/10.1109/URKE.2011.6007796

Sharda, R., Delen, D., & Turban, E. (2018). *Business intelligence, analytics, and data science: A managerial perspective* (Fourth edition). Pearson.

Ünvan, Y. A. (2020). Market basket analysis with association rules. *Communications in Statistics - Theory and Methods*, 1–14. https://doi.org/10.1080/03610926.2020.1716255

Vahidi Roodpishi, M., & Aghajan Nashtaei, R. (2015). Market basket analysis in insurance industry. *Management Science Letters*, *5*(4), 393–400. https://doi.org/10.5267/j.msl.2015.2.004

Zekić-Sušac, M., & Has, A. (2015). Discovering market basket patterns using hierarchical association rules. *Croatian Operational Research Review*, *6*(2), 475–487. https://doi.org/10.17535/crorr.2015.0036

Biographies

Ivan Diryana is a faculty member at Bina Nusantara University, graduated from doctoral of business management, he has an experience in industry for 2 years then focusing in education field as a lecturer. He also have a business in culinary area for more than eight years. Interest area of research and lecturing in entrepreneurship, marketing, business, and management.

Ronny Samsul Bahri is a faculty member of entrepreneurship department at Bina Nusantara University. Currently he pursuing his doctoral at Parahyangan University, Bandung. His research interest including entrepreneurship, data analytics, business, and management.

Iston Dwija Utama is a faculty member of entrepreneurship department at Bina Nusantara University. He graduated from Master in Business Administration program from Bandung Institute of Technology, passionate in entrepreneurship, marketing, and management field. Experienced in the industry field for more than 8 years and having own business in culinary. Besides teaching, he is also active in a professional organization that concern in SMEs namely LUNAS (Layanan UMKM Naik Kelas) as a business mentor and HIPMI Kab Bandung.

Chyntia Ika Ratnapuri is a faculty member of entrepreneurship department at Bina Nusantara University and Business Incubator Section head at Binus business incubator Bandung. She earned bachelor degree in management UNPAR and master in management business administration SBM ITB. Chyntia Ika's research interest include entrepreneurship, design creative and innovative thinking, and also business model in creative industry.