MARKET BASKET ANALYSIS: ANALYSIS OF ASSOCIATION RULES IN DIFFERENT TIME PERIODS

BY

TENG WEN FOONG

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

BACHELOR OF COMPUTER SCIENCE (HONOURS)

Faculty of Information and Communication Technology

(Kampar Campus)

JAN 2024

UNIVERSITI TUNKU ABDUL RAHMAN

REPORT STATUS DECLARATION FORM

Title: Market Basket Analysis: Analysis of Association Rules in Different Time Periods

Academic Session: JAN 2024

I TENG WEN FOONG

(CAPITAL LETTER)

declare that I allow this Final Year Project Report to be kept in
Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.

Verified by,

(Author's signature)

(Supervisor's signature)

Address:
No. 2, Lorong Sejahtera 1,
Taman Sejahtera, 36700
Langkap, Perak

Dr Kh'ng Xin Yi
Supervisor's name

Date: 25/4/2024

Date: 25/4/2024

FACULTY OF _INFORMATION AND COMMUNICATION TECHNOLOGY (KAMPAR CAMPUS)_

UNIVERSITI TUNKU ABDUL RAHMAN

Date: ___25/4/2024____

SUBMISSION OF FINAL YEAR PROJECT

It is hereby certified that _____***Teng Wen Foong***___ (ID No: __***20ACB01778***__) has completed this final year project entitled "__*Market Basket Analysis: Analysis of Association Rule in Different Time Periods* " under the supervision of _____Dr Kh'ng Xin Yi_____ (Supervisor) from the Department of ___Computer Science___, Faculty/Institute* of __Information and Communication Technology (Kampar Campus)__ , and ____ (Co-Supervisor)* from the Department of _____, Faculty/Institute* of _____.

I understand that University will upload softcopy of my final year project in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_____
(*TENG WEN FOONG*)

*Delete whichever not applicable

DECLARATION OF ORIGINALITY

I declare that this report entitled "**MARKET BASKET ANALYSIS: ANALYSIS OF ASSOCIATION RULES IN DIFFERENT TIME PERIODS**" is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature : _____

Name : Teng Wen Foong

Date : 25/4/2024

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and appreciation to my supervisor, Dr Kh'ng Xin Yi for invaluable guidance and unwavering support throughout my final year project. Insightful feedback, constructive critiques, and thoughtful suggestions from Dr Kh'ng have consistently propelled me forward, helping me navigate the complexities of my project with confidence. Again, much appreciation for the exceptional mentorship, encouragement, and support. I look forward to continuing to learn and grow under her guidance as I move forward in my academic and professional endeavors.

ABSTRACT

Transaction data is a set of records associated with the sales-purchase activities of a business. As a result of digitalization, the transaction data stored in databases have been used by the analyst to extract valuable information. The most widely used method to discover information from transaction data is market basket analysis. Market basket analysis is a technique that determines the relationship between items and the relationship is known as the association rule. The association rules could be further studied to identify commonly purchased items and gain a deeper understanding of customer purchasing behaviours. In this study, performing market basket analysis is not the only objective but the variation of association rules over time will also be focused on. Therefore, market basket analysis will be conducted in different time periods to study the regular cyclic variation of association rules over time and determine the interaction between association rules and time. After analysing the trend of the customers purchasing behaviours from the association rules, the interesting association relationships between large quantities of business transaction data could assist the businesses to design marketing strategies, promote cross-selling, and other business decision-making to boost their profit. An optimal period to conduct market basket analysis will be recommended to ensure the effectiveness of the association rules is maximized all the time and allow the businesses to adjust their marketing strategies from time to time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Problem Statement and Motivation

Most of the business owners have ignored the need of tracking the changes of association rules over time and this circumstance might cause them to miss out on the adaptions needed to be made in order to retain the effectiveness of the association rules. The association rules could be affected by the case of certain existing items are removed or some new items are added to the inventory. This circumstance may lead to customers start purchasing the new items which will result in a new association rule or the existing association rule might lose its effectiveness due to the item no longer existing. Nevertheless, customer preferences might change due to introduction of new trend. This could happen when some of the product companies put a lot of effort to advertise their products for better exposure to the public. For instance, release of a scientific report which stated that banana milkshake will benefit human body in many ways will attract more customers to buy banana and milk together. As a result, customers may follow the new trend and change their buying behavior. Additionally, external factors such as weather will also affect the effectiveness of market basket analysis. Demand of seasonal products like clothing changes over time. For instance, thick and warm clothing such as scarves, gloves, and earmuffs will have higher demand during the winter season. Moreover, competition between businesses in market will also affect the market basket analysis over time. This is due to new competitors joining the market may offer several promotions and discounts for certain product combinations to attract customers and boost their sales. For instance, customers who buy milk and egg together will receive a discount of 20%. Eventually, it will lead to a new association rule and it might not be found if market basket analysis is not conducted periodically. Likewise, external events such as global pandemic could be recognised as one of the key factors that affect market basket analysis over time as well. For instance, the global pandemic COVID-19 which happened in these few years causes the high demand of items that protect human body from the virus. For instance, the frequency of item combination like medical mask and hand sanitizer has increased dramatically compared to other groceries during pandemic. In conclusion, there are several factors require businesses to conduct market basket analysis periodically to guarantee the elimination of insignificant association rules and the discovery of new potential association rules.

By conducting market basket analysis and analysing association rules in different time periods, businesses and retailers are allowed to track changes of customer behavior over time. Subsequently, businesses that are facing a regression of their sales will benefit from the market basket analysis. Additionally, it will help businesses to adapt their marketing strategies by focusing on certain products instead of retaining the same old strategies. Moreover, conducting market basket analysis in different time periods will help the businesses to discover a new trend or opportunities to increase their revenues. Assisting businesses to survive in the market is not only the motivation for conducting the analysis, customers will also benefit from it. For instance, businesses may place items that have association rules together to help customers with reduced mobility to find them without needing to move around the store to look for them. In conclusion, market basket analysis is a result of the rise of information technology and it could benefit society from both customer and businesses perspectives.

1.2 Objectives

1.2.1 To conduct market basket analysis in different time periods.

Unlike common market basket analysis, the project will conduct a more specific market basket analysis with an additional factor which is the time period. Afterward, an analysis of transaction data from different time periods such as quarters, months, and biweeks, will be carried out using market basket analysis. The project will be initiated by conducting a market basket analysis using a dataset consisting of transactions in different time periods to determine the association rules. When conducting MBA, the execution time will be observed to assess the efficiency and performance of the analysis process. The execution time can vary depending on various factors, such as the size of the dataset in different time periods, the complexity of the analysis algorithm and the computational resources available for MBA.

1.2.2 To analyze the fluctuation of the customer purchasing behaviors with the changes in time.

By comparing the statistics of association rules over different time periods, a deeper understanding and better insight into why and how customer behavior changes over time will be obtained. The parameters of the association rules such as confidence, lift, and support will be visualized in the forms of graphs, charts, or plots for the observation

of which of the values change over time. The hypothesis could be proven by comparing the real transaction data to check if the fluctuation of the transaction frequency. Next, the variation of customer purchasing behaviors with the changes in time will be analysed to understand how they change and identify if any external factors that may be influencing these changes. Customer purchasing behaviors could change in several ways such as shifting of product preferences, fluctuation of purchasing amount, and so on.

1.2.3 To provide actionable recommendations for optimal time period for market basket analysis.

Product recommendations will be provided to the businesses, along with the justifications, for tracking and analyzing the changes in customer behavior more effectively. The insights gained from the project will benefit businesses to adjust their marketing strategies from time to time (e.g. product placement, and pricing of the products) after understanding how the association rules variate over time. Nevertheless, recommendations for cross-selling and up-selling could be made by considering the external factors affecting the fluctuation of customer purchasing behaviors such as the frequency of customer purchases, length of time between purchases, and the seasonality of product demand. An optimal time period for conducting market basket analysis will guide the businesses to maximize the effectiveness of their marketing strategies from time to time. Eventually, the revenue of the businesses and customer satisfaction will be improved.

1.3 Project Scope and Direction

The project will focus on the analysis of association rules in different time periods using market basket analysis. Market basket analysis will be conducted using a dataset of customer transactions which are divided into different time periods such as quarterly, monthly, and biweekly. First of all, the data collected from the retail store will undergo a data preprocessing process which involves data cleaning and preparation of the data for analysis later on. In this phase, the data will be checked for completeness and consistency to guarantee quality of data. Any missing values will be input, and the outliers will be removed. Subsequently, market basket analysis will be performed using the prepared data to discover association rules in different time periods. The discovered

association rules will then undergo a filtering process which will disqualify those that do not meet the minimum requirement in terms of the value of the important parameters in market basket analysis. The association rules which fulfill the requirements will be visualized in the form of graphs, charts, or tables using their parameters such as lift, support, and confidence. By presenting the data visualization, changes of customer behavior could be observed based on the variation of the values in different time periods. Next, data interpretation will focus on the time period with the most significant variation of values to figure out the factors affecting the association rules. Further analysis will be carried out to investigate the best time period to conduct market basket analysis in order to track changes of association rules. Eventually, a recommendation with detailed justification will be provided for the businesses to refer for improving their product sales and boosting their revenues. However, there might be some constraints when carrying out the project. For instance, the large size of the dataset which included transaction data in different time periods might be constrained by an available source such as time. The deliverables of the project are association rules for different time periods, analysis of trends and patterns in association rules over time, analysis of factors affecting the variation of association rules, and recommendations to conduct market basket analysis in terms of the time period to adjust marketing strategies and improve sales.

1.4 Contributions

The main beneficiary of the project will be the businesses. The market basket analysis to be conducted in different time periods will assist them to make more strategic business decisions. First of all, businesses will gain better insight into the placement of their products in stores. The association rules from market basket analysis help the businesses to identify the popular combination of products and place them together. This will allow customers to find their desired products more conveniently and encourage them to make additional purchase to boost sales. Moreover, a new pricing strategy may be discovered as a result of the project. For instance, the businesses could offer a bundled promotion for the products involved in the association rules to attract customers to buy the products together. By conducting the analysis, businesses could improve their inventory management by stocking up on products that are frequently purchased together, reducing the possibility of stockouts to maximize profits. Last but

not least, businesses will be able to provide a better customer experience by recommending products that are frequently purchased together, based on the association rules. The customers might accept the recommended products that they did not initially think of, eventually boosting the product sales.

1.5 Report Organization

The report is organized into five distinct chapters, which are introduction, literature review, proposed method / approach, preliminary work, and conclusion. The first chapter outlines the project's aims, importance, and context, which could be the foundation for the upcoming chapters by determining research inquiries and objectives. The second chapter, literature review will discuss the related work and past researches, then study further based on the works done by others. This section will address limitations of each work and the limitations will be taken into account in current project to resolve them. Subsequently, the following chapter, proposed method / approach will explain the methodology in detailed, elucidating how the project harnesses market basket analysis across various time spans to pinpoint the most efficient period for association rule mining. This chapter would also cover the algorithms to be used, tools, and data processing techniques deployed. To offer practical insights, the chapter of preliminary work investigates the operations that will be performed in the project, including data gathering, preprocessing, and initial findings, illustrating progress and laying the groundwork for the conclusive chapter. Last but not least, the conclusion chapter summarizes project outcomes, underscoring pivotal discoveries, trends, and implications arising from the association rule mining analysis. It also reflects on the importance of the optimal time period findings and explores potential avenues for future research and practical implementation.

CHAPTER 2 LITERATURE REVIEW

2.1 Market Basket Analysis

Market basket analysis is known as a data mining technique that analyses the patterns of co-occurrence of certain products and investigates the relationship and strength of the link between the products that have been purchased together by the customers. The main objective of market basket analysis is to identify the purchasing behavior of the customers in order to help retailers to understand better which are the popular products and try to get more stocks of those certain items in the inventory. In addition, retailers could find out the best strategic location to display their products by placing the products that are being bought together frequently in the same area. By having a more effective product placement, retailers could boost their sales by a significant amount. The concept of market basket analysis could be interpreted using an example from daily life. When people buy a loaf of bread, they tend to buy fruit jam as well. The relationship between bread and jam is recognised as an if-then conditional statement which implies that if bread is bought, then the jam will be bought with bread. Market basket analysis can be categorised into three types which are descriptive market basket analysis, predictive market basket analysis, and differential market basket analysis. This type of market basket analysis is done by analysing historical data and identifying the relationship between the products using statistical techniques. In fact, there will be no prediction on the future trend of the buying behavior of customers will be made and this kind of analysis is referred as an unsupervised technique. Next, predictive market basket analysis works reversely with descriptive market basket analysis. It can be interpreted by its literal meaning "predictive", analysis of historical and current existing data will be carried out before predicting the future trend. This kind of analysis makes use of supervised machine learning models such as classification and regression to obtain accurate predictions based on the provided data. However, this approach is less commonly to be used when compared to descriptive market basket analysis. The last type of market basket analysis is differential market basket analysis. This type of analysis could help retailers to understand why customers will purchase a product from a certain store even though there are other stores with similar product prices and quality. In this analysis, the customer decisions will be studied based on several factors such as

seasons, time variations, and even delivery time and user experience for that online shopping on e-commerce platforms [1].

2.1.1 Association Rule

Association rule is the core concept of market basket analysis which focused on determining insignificant and hidden relationships between two items in a large data set. The data set to be analysed consists of customer transaction activity and one of the sources for the data is supermarkets. The transaction activity data will list the items that have been bought by customers in a single transaction. Analogously, association rule is similar to market basket analysis such that association rule is also categorised into four types which are multi-relational association rules, generalized association rule, interval information association rule, and qualitative association rule. Multi-relational association rule (MRAR) is derived from multiple different relational databases. Each rule under this type will contain an entity that represents the indirect relationships between entities by having different relationships. The generalized association rule is the most commonly used association rule in data mining to get a fundamental idea about the hidden patterns in data [2]. Kristiana et al. in 2020 have highlighted that the idea of the association rule is to identify all the possible connections between items and select the most frequent combination of items that will most likely be indicators of dependence [3]. The statement of association rules is developed in two stages as proposed by Rusawati, Gufroni, and Rianto in 2018 [4]. The initial stage to develop association rules is analysing patterns of high frequency. In this phase, the combination of items that meet the minimum support value that has been determined will be figured out. The association rule is supported by a mathematical model. In the application of the model, different products are recognised as objects in a set, I = {i1, i2, i3, i4, … …, in}. Each transaction is expressed as T and it corresponds to objects in the set I, which could be defined as $T \subset I$. Assuming that A and B are the object sets that fulfill the conditions $A \subset I$, $B \subset I$, and $A \cap B = \emptyset$, then the circumstance has achieved the association rule which can be expressed as $A \Rightarrow B$ [5]. Based on the definition of the association rule, A is the antecedent (if) and B is recognised as the consequent (then). After determining all the high-frequency item sets, the development of the association rule proceeds to the next stage which is association rules establishment. The association rule, $A \Rightarrow B$ is only considered as confident after it meets the minimum requirement of

confidence value. The terms, “support” and “confidence” will be further discussed in the following subchapters.

2.1.2 Support

Support is one of the parameters that defined the market basket analysis. Support is the percentage of the combination of certain items in a large amount of data. The support value shows how frequently the combination of items appears in transaction data. In fact, association rules with a higher support value will be preferred due to they are more likely to show the pattern in large transaction data. The support value is crucial for the first stage of developing association rules which have been mentioned earlier. The high-frequency analysis will only select the combination that reaches the minimum support value. Nevertheless, support could also be referred to as a probability. For instance, if bread is having a supporting value of 20 out of 100 transactions, then it indicates that the probability of the bread being sold is 0.2. In terms of mathematics, the supporting value could be obtained through the formula of dividing the amount of the transaction that contains the specified item by total transactions and multiplying the value by a hundred percent. The mathematical formula can be expressed as the following shows:

$$S = \frac{\sum(Ta + Tb)}{\sum(T)}$$

$$Support = \frac{The\ number\ of\ transaction\ that\ contains\ antecedent\ and\ consequent}{Total\ transaction}$$

where:

S = Support value

$\sum (Ta + T\text{b})$ = Transactions that contains item A (antecedent) and item B (consequent)

$\sum(T)$ = Total transactions

2.1.3 Confidence

Confidence is another important measure for market basket analysis. It works similarly to support but it is more focused on conditional probability. The confidence value emphasises the strength of relationships between two certain items in association rules. For instance, the analyst could obtain the confidence value by studying how frequent

the occurrence of the combination of bread and jam is among the total frequency of bread. Subsequently, the confidence value will be applied in the second stage of developing association rules which are establishing association rules by disqualifying the combinations that do not meet minimum confidence requirements. In 2018, Sagin and Ayvaz have developed association rules mining with market basket analysis for a large hardware company operating in the retailed sector [6]. The project aimed to identify product groups that tended to be sold together and eventually they figured out the rule of {Chisels and Cutters, Bits Tips Adapters} → {Drill Bits} with a 64% confidence value which is considered high. The company was totally unaware of the hidden rule before the project is being carried out. In terms of mathematics, confidence value has applied conditional probability and it is expressed as the following formula:

$$C = \frac{\sum(Ta + Tb)}{\sum(Ta)}$$

$$Confidence = P\,(A \mid B)$$

$$= \frac{The\ number\ of\ transaction\ that\ contains\ item\ A\ and\ B}{The\ numebr\ of\ transaction\ that\ constains\ item\ A}$$

where:

C = Confidence value

$\sum\ (Ta + T\text{b})$ = Transactions that contains item A (antecedent) and item B (consequent)

$\sum(T)$ = Total transaction that contains A

2.1.4 Lift

In market basket analysis, a lift is a metric that is used to measure the strength of the relationship between two items as a combination. Lift is calculated by dividing the confidence percent by the support percent. A lift value of 1 indicates that the association rule does not have any effect on the outcome which means that there is completely no association between the two items. If the value of lift is greater than 1, it implies a positive association which shows the association rule enhances the chances of the outcome, while a lift value lower than 1 corresponds to the situation of a negative association. A higher lift value indicates a higher possibility of the presence of one item in a transaction coming with another item in the same transaction. The mathematical formula for calculating lift is shown below:

$$Lift(A \Rightarrow B) = \frac{Confidence\ (A \Rightarrow B)}{Support\ (B)}$$

2.2 Apriori Algorithm

The Apriori algorithm is the first and most widely used data mining technique that was proposed by R. Agrawal and R. Srikant in 1994 for frequent itemset mining [7]. It is a bottom-up approach that uses a level-wise search for identifying the frequent item sets. There are two main steps in the Apriori algorithm which are the join step and the prune step. The join step generates (k+1)-itemsets from k-itemsets by joining each item with itself while the prune step determines the frequent itemsets that meet minimum support and disqualify those itemsets that are recognised as infrequent. In the first iteration of the algorithm, a 1-itemsets candidate is found by scanning the database in which the occurrence of the itemset is satisfying the minimum support threshold. Next, 2-itemsets that meet minimum support are determined using previous frequent 1-itemsets and this process will continue until frequent k-itemsets are found [8]. In 2021, Nurmayanti et.al. applied the Apriori algorithm to determine association rules between several sales attributes based on the transaction data of outdoor sales of goods [9]. As a result, they have found that the transaction patterns in the sale of outdoor goods that are generated with the algorithm are as many as 10 rules and the strength of the rules is supported by 0.296 minimum support, 0.774 confidence, and 1.49 lift value. They could assume that customers who buy Portuguese stove items will also be likely to purchase portable gas items. The pseudocode for the Apriori algorithm is shown in the figure below [10]:

```
procedureApriori (T, minSupport) { //T is the database and min-Support
is the minimum support
Ck: Candidate itemset of size k
Lk: frequent itemset of size k
L1= {frequent items};
for(k= 1; Lk !=∅ ; k++) do begin
Ck+1= candidates generated from Lk;
for each transaction t in database do{
increment the count of all candidates in Ck+1 that are contained in t
Lk+1= candidates in Ck+1 with min_support
}end
return UkLk;
```

Figure 2.2.1. Pseudo code for Apriori algorithm [10]

In 2018, Ayvaz and Sagin conducted market basket analysis to determine association rules for a large hardware company by using Apriori algorithm [6]. As a result, they

successfully develop the first rule and it is shown in the figure below. They found that there are 431 out of 667 transactions containing the combination of {Chisels and Cutters, Bits Tips Adapters} also contain product {Drill Bits}. This rule indicates that customers who purchase the group products {Chisels and Cutters} and {Bits Tips Adapters} are having the 64% probability of buying {Drill Bits} as well.

Antecedent	*Consequent*	*Confidence*	*Lift*
Chisels and Cutters, Bits Tip Adapters (677)	Drill Bits (431)	%64	3
Chisels and Cutters, Punch Group (366)	Drill Bits(233)	%64	3
Chisels and Cutters, Wire Brush (454)	Drill Bits (271)	%60	2.81
Chisels and Cutters (2370)	Drill Bits (1361)	%57	2.70
Meters, Nut Adapters (411)	Drill Bits (235)	%57	2.69
Diamond Saw, Chisels and Cutters (563)	Drill Bits (316)	%56	2.64
Meters, Nut Adapters (481)	Bits Tip Adapters (251)	%52	3.33
Nut Adapters, Drill Bits (962)	Bits Tip Adapters (497)	%52	3.29
Wire Brush, Nut Adapters (618)	Bits Tip Adapters (305)	%49	3.15
Bits Tip Adapters, Files(505)	Drill Bits (242)	%48	2.26

Figure 2.2.2. Product affinity results [6]

2.3 Comparison between Apriori and FP-Growth Algorithm

The Frequent Pattern Growth (FP Growth) algorithm is an alternative data mining algorithm for market basket analysis. The algorithm could be recognised as an improved version of the Apriori algorithm since it overcomes the flaws of the Apriori algorithm by eliminating the join and prune steps. The first step of the algorithm is identical to the first step of Apriori, which is scanning the database to obtain the frequent itemsets. The non-frequent itemsets are filtered out based on the minimum support and the remaining itemsets are listed in order of size. Next, a tree with only a root is created and the frequent itemsets are added as nodes into the tree one by one. The construction of an FP-Tree is considered complete once all the frequent itemsets have been added to the tree. Now, it is much faster to get information on the most frequent itemsets by traversing the tree [11]. There are several differences between the

Apriori algorithm and the FP Growth algorithm. Regarding the generation of frequent patterns, Apriori obtains itemsets by pairing such as single itemset, double itemset, and triple itemset, while FP Growth constructs an FP-Tree. In addition, Apriori uses frequent itemsets to extend frequent subsets one item at a time whereas FP Growth constructs conditional FP-Tree for every item in the database. In the aspect of time consumption, Apriori takes longer time compared to FP Growth since it scans the database at each step. In contrast, FP Growth scans the database only once at the beginning so it saves more time. In the context of space, Apriori saves a converted version of the database in memory, while FP Growth saves the set of conditional FP-tree for each item in the memory. Last but not least, Apriori applies bread-first search to search for items while FP Growth uses depth-first search due to its data structure [12]. In 2021, Aldino et al. carried out a market basket analysis to determine customer purchasing patterns using FP Growth and Apriori algorithm [13]. They found that FP Growth had a faster running time of 6 seconds which is 5 times shorter than Apriori which required 30 seconds. The last conclusion drawn from their studies is that FP Growth generates more than 19 association rules while Apriori only produces 6 rules. Hossain et.al conducted market basket analysis based on a bakery shop dataset using Apriori and FP Growth algorithms respectively [14]. They found that the required time for the FP Growth algorithm is much less than that for Apriori algorithm. The graph of comparison of execution time is shown as below:

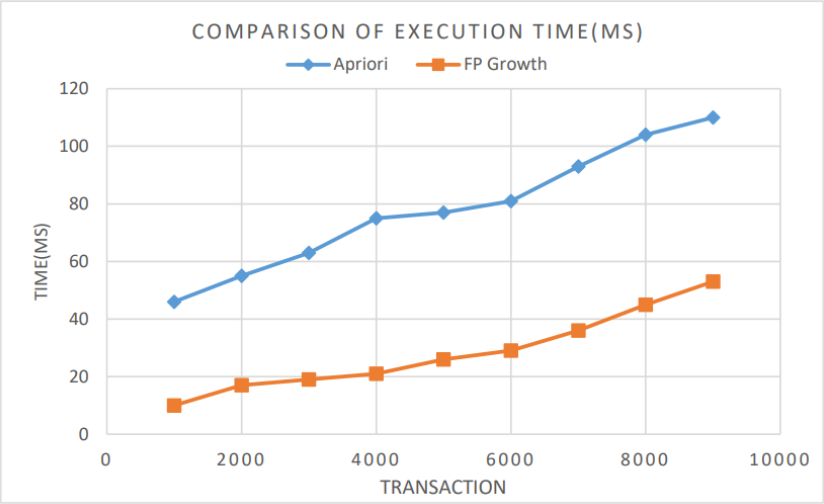


Figure 2.3.1. Comparison of execution time of Apriori and FP Growth algorithms [14]

2.4 Time Variations of the Association Rule

Time variations of the association rule are referred to as the possible changes in the association rule over time. Customer behaviour may be different due to several factors such as market trends, product availability, and even changes in environmental issues. For instance, people are buying warm and thick clothes to wear during winter and looking for light and thin clothes in summer. Consequently, it is crucial to detect any significant changes in association rules over time to monitor customer behaviour and adjust the marketing strategy accordingly. Computing association rules based on monthly transaction data will help the analyst to recognise seasonal variation where some of the rules indicate a pattern in approximately the same month each year. Analogously, different time variations such as hourly, daily, weekly, and even annually might dig out some hidden association rules. In 2019, Ansari did a study that analyses the trend of association rules in different time periods. He worked on transactional data of a Belgian retail company and analysed the results that will assist the company to adjust time period-specific marketing strategies to boost the revenue of the company [15]. He also stated that most of the retail stores are prioritizing what the customers will purchase but they have ignored another crucial factor that is affecting their behaviors. In fact, they should not only focus on the popular combinations of items but also pay attention to the fact about when they buy them. Ansari also presented the statement of the need of focusing on the problem that shows regular variations in association rules over time. In his studies, he applied 3 months periods to discover top association rules and compare them at different time periods which will allow the company to perform up-selling and cross-selling in different time lags and maximise the profit over time. In 2011, Papavasileiou and Tsadiras presented how variations of association rules occur over time by using market basket analysis [16]. They studied the connection between variability of association rules with changes in purchasing habit over time. Annual data with 12 subsets of data corresponding to each month of a year was collected for analysis purpose and a significant change was observed during first two months. The variation of association rules was determined by three key factors: lift, support, and confidence. Based on the figure below, the table shows that rule 5, 6, and 7 were having a massive variation in terms of confidence. Rule 5 was having a growing confidence while confidence of rules 6 and 7 were decreasing.

		Month 1 (Sept. 2008)			*Month 2 (Oct. 2008)*		
A/A	*Rules*	L	S %	C %	L	S %	C %
1	Yoghurt – Bread > Fruits	1.83	1.27	86.84	1.72	1.49	81.02
2	Toast- Cereals> Fresh Milk	1.20	2.26	71.89	1.23	2.52	73.81
3	Fruits – Sweet Biscuits > Vegetables	1.73	2.63	64.70	1.79	2.44	61.77
4	Fruits – Pasta > Fresh Milk - Vegetables	1.83	1.87	40.70	1.84	1.69	36.44
5	Fresh Milk – Beers > Cola	1.86	1.27	40.60	11.2	1.12	66.42
6	Cola – Salad (Not Packed, Packed) > Vegetables	12.8	1.02	79.50	1.14	1.41	68.28
7	Chicken Not Packed> Fruits	14.1	1.57	53.49	1.17	1.01	30.55
.	……..	…	…	…	…	…	….
.	……..	…	…	…	…	…	….
1803	Fresh Milk – Beef Not Packed> Bread	1.63	1.28	33.67	1.73	1.18	34.36
		….. .			*Month 12 (Aug. 2009)*		
		L	S%	C%	L	S%	C%
		…	…	…	1.85	1.26	85.95
		…	…	…	1.38	2.58	79.41
		…	…	….	1.56	2.70	62.62
		…	…	….	1.61	1.53	36.42
		…	…	…	1.69	1.71	37.79
		…	…	…	15.2	1.19	99.34
		…	…	….	1.16	1.26	32.78
		…	…	….	…	…	….
		…	…	….	…	…	….
		…	…	….	1.58	1.34	31.12

Figure 2.4.1. Monthly Development of Association Rules Measurement Indicators [16].

Alfiqra and Khasanah in 2020 developed a model to implement market basket analysis based on Overall Variability of Association Rule (OCVR) on product marketing strategy [17]. The OCVR is a metric used in market basket analysis to evaluate the stability or variability of association rules over time. A higher OCVR indicates greater variability and suggests that the association between the items change significantly across time. Their study used monthly transaction data of a retail store in Yogyakarta to produce rules for four periods, which corresponded to weeks. As a result, they found 59 rules for the first period, 48 rules for the second period, 54 rules for the third period, and 58 rules for the last period. Based on the following figures, it can be observed that rules obtained from each period had a slight difference while the attributes in terms of support, confidence, and lift were changing over the four periods. This indicated that customer behavior changed over time, hence, marketing strategy should be adjusted from time to time to maximize the benefits from market basket analysis. The formula for computing OCVR is shown below:

$$CV = \frac{s}{\bar{x}}$$

where:

CV = Variability index

s = Standard deviation

$\bar{x}$ = Mean value

$$OCVR = \frac{CVL + CVC}{2}$$

where:

OCVR = Overall variability

CVL = Index variability lift

CVC = Index variability confidence

Table 2.4.1. Rules for 1st period [17]

No	Rule			Support	Confidence	LR	Count
1	{Liquid soap}	→	{Bar Soap}	0.00109	0.62500	74.68	15
2	{Noodle Y}	→	{Noodle X}	0.00116	0.57143	7.14	16
3	{Milk cleanser}	→	{Face tonic}	0.00437	0.52174	60.25	60
4	{Noodle Z}	→	{Noodle X}	0.00138	0.51351	6.41	19
5	{Face tonic}	→	{Milk cleanser}	0.00437	0.50420	60.25	60
. . .	. . . .	→	. . .	. . .	. . .	. . .	. . .
59	{Margarine}	→	{Noodle X}	0.00167	0.20000	2.49836	23

Table 2.4.2. Rules for 2nd period [17]

No	Rule			Support	Confidence	LR	Count
1	{Noodle Y}	→	{Noodle X}	0.00104	0.59259	8.51	16
2	{Face tonic}	→	{Milk cleanser}	0.00267	0.53247	87.06	41
3	{Noodle A}	→	{Noodle X}	0.00111	0.50000	7.18	17
4	{Chili sauce}	→	{Noodle X}	0.00124	0.47500	6.82	19
5	{Milk cleanser}	→	{Face tonic}	0.00267	0.43617	87.06	41
. . .	. . . .	→	. . .	. . .	. . .	. . .	. . .
48	{Biscuit}	→	{Wafer}	0.00156	0.20000	3.06	24

Table 2.4.3. Rules for 3rd period [17]

No	Rule			Support	Confidence	LR	Count
1	{Noodle B}	→	{Noodle X}	0.00123	0.53125	7.57	17
2	{Face tonic}	→	{Milk cleanser}	0.00325	0.46392	61.13	45
3	{Biscuit}	→	{Wafer}	0.00116	0.45714	6.42	16
4	{Liquid soap}	→	{Bar Soap }	0.00101	0.45161	71.01	14
5	{Milk cleanser}	→	{Face tonic}	0.00325	0.42857	61.13	45
. . .	. . . .	→	. . .	. . .	. . .	. . .	. . .
54	{Dish wash}	→	{Noodle X}	0.00202	0.20000	2.85	28

Table 2.4.4. Rules for 4th period [17]

No	**Rule**			**Support**	**Confidence**	**LR**	**Count**
1	{Face tonic}	→	{Milk cleanser}	0.003774	0.56000	72.88	56
2	{Noodle C}		{Noodle X}	0.001348	0.55556	7.40	20
3	{Liquid soap}	→	{Bar Soap}	0.001146	0.53125	85.68	17
4	{Cat food A}		{Cat food B}	0.001146	0.50000	176.63	17
5	{Milk cleanser}	→	{Face tonic}	0.003774	0.49123	72.88	56
. . .	. . . .		. . .	. . .	. . .	. . .	. . .
58	{Soy sauce}	→	{Noodle X}	0.001078	0.20779	2.77	16

CHAPTER 3 SYSTEM MODEL

3.1 Design Specification

Within this section, an intricate layout and the essential technical prerequisites of the project's resolution will be expounded. Comprehensive direction for crafting the solution such as encompassing the solution's architecture, constituents, data progression, and other intricate technical facets will be provided. This meticulous guidance aims to preempt any potential misinterpretations that could arise from ambiguities in requirements. Moreover, the design specification serves to guarantee that project advancement remains in alignment with its objectives, design principles, coding conventions, and other established norms. The subchapter spans three primary aspects: methodologies and general work procedures, resource requirements, and system performance definition.

3.1.1 Methodologies and general work procedures

The project methodology will involve the utilization of both the Apriori and FP-Growth algorithms for conducting association rule mining. Although these algorithms yield identical association rule outputs, however, their performance varies in terms of processing time. Therefore, a comparative analysis of their processing times will be conducted to observe the impact of this distinction. The Apriori algorithm is a traditional association rule mining technique which its main objective is to mine frequent itemsets within transactional datasets. It operates iteratively by generating candidate itemsets and pruning those that do not meet a defined minimum support threshold. However, the algorithm's repeated data scans and candidate generation can render it computationally intensive for sizable datasets.

In contrast, the FP-Growth algorithm is a sophisticated association rule mining technique designed to overcome Apriori's limitations such as its multiple dataset scans and candidate generation stages. FP-Growth constructs a compact data structure named the FP-tree which encapsulates item frequency and relationships within the dataset. This construction obviates the necessity for candidate generation. Instead, the algorithm recursively mines frequent itemsets directly from the FP-tree which is considered curtailing computational overhead. FP-Growth excels with datasets featuring numerous transactions and items due to it necessitates fewer data passes and eliminates the need

to generate an extensive array of candidates. This attribute renders FP-Growth an influential technique for scalable association rule mining.

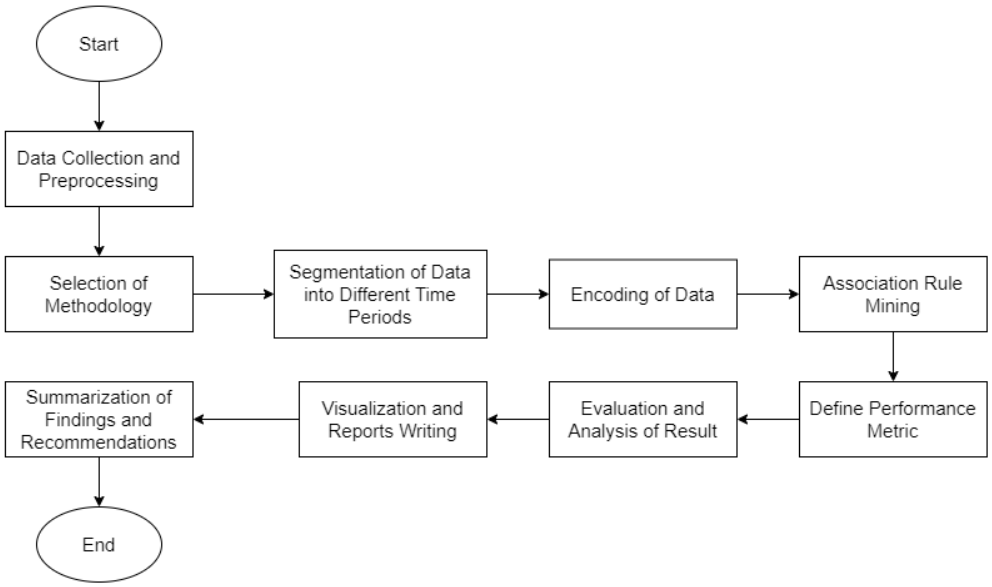


Figure 3.1.1. System flow diagram

The initial step of the project involves initiating data collection and preprocessing procedures. The collection phase entails procuring transactional data essential for conducting market basket analysis. This dataset will encompass records of customer purchases and unveils valuable insights into customer buying behaviors. Possible sources of this data include point-of-sale systems, e-commerce platforms, customer databases, and loyalty programs. To be considered valid, transactional data must encompass key attributes such as invoice number, item stock code, item description, quantity, invoice date, and unit price. The preprocessing phase ensues once the raw data is amassed. To facilitate subsequent association rule mining, data transformation into an appropriate format is requisite and integration becomes necessary if the data originates from diverse sources. Essential data-cleaning procedures are undertaken to expunge duplicates, and irrelevant entries, and rectify missing values through imputation or elimination.

The project then proceeds to segment the data into distinct time periods to align with the project's objectives. The specified intervals consist of quarterly, monthly, and biweekly spans aimed at extracting meaningful insights. The dataset is organized chronologically based on the invoice date timestamps. Due to memory constraints, only

the three consecutive months displaying the highest transaction frequency will be selected. Transactions within these months are grouped, and transactions sharing the same invoice number are considered a single transaction. Following the preparation of monthly transactions, biweekly transactions are derived from the preprocessed monthly data. The prepared data then undergoes encoding. Data encoding is a critical preprocessing step that transforms categorical data into a numerical format, amenable to algorithmic analysis. For market basket analysis, this involves converting item names into binary representations indicating item occurrences in transactions. In this project, one-hot encoding will be employed to ensure each unique item is allocated a binary column in the dataset. Transactions form rows in such a way that "1" denotes an item's presence and "0" its absence. This conversion enhances data processing efficiency for algorithms like Apriori and facilitates the discovery of frequent item sets and association rules.

With the encoded data prepared, the algorithms can be employed for association rule mining. This method unveils connections between items frequently co-occurring in transactions and reveals insights into customer buying patterns. Frequent item sets are identified, and association rules are deduced from these sets to depict conditional relationships between items. Each rule comprises an antecedent (premise) and a consequent (outcome) item set and they are accompanied by metrics such as support, confidence, and lift that denote the rule's strength and significance. Various metrics are utilized to gauge rule quality and relevance. Support, lift, and confidence metrics will be used to assess the reliability of item relationships, hence, guiding informed decision-making. A judicious minimum support threshold is defined to sieve out insignificant rules.

Subsequently, an evaluation and analysis of the association rules is conducted to extract meaningful insights. This entails scrutinizing the discovered rules and singling out high-confidence and high-lift rules that illuminate customer behavior. A cross-comparison of rules across different months and biweekly periods is executed to pinpoint trends and shifts in customer preferences. Visual aids such as bar charts serve to effectively communicate patterns and trends embedded in the identified rules. The ensuing report employs succinct yet informative narratives to expound upon rule significance, and spotlight significant findings, temporal changes, and potential implications for business strategies. The report's coherence ensures comprehension

among both technical and non-technical stakeholders, empowering them to make business-influencing decisions rooted in mined association rules.

The findings are subsequently distilled into concise, actionable takeaways, summarizing complex insights. They synthesize pivotal patterns, trends, and variations identified through rule mining, accentuating their potential impact on business objectives. In this analysis, insightful recommendations are formulated and suggestions in terms of strategies for marketing campaigns, inventory management, or customer segmentation will be recommended. This summary streamlines comprehension of the analysis results and facilitates strategic decision-making guided by data-driven insights. This approach optimizes resource allocation, augments business performance, and enhances overall operational efficiency.

3.1.2 Resource requirements

Resource requirements pertain to the distinct tools, materials, technologies, equipment, and personnel indispensable for the proficient execution of a project or undertaking. These resources play a pivotal role in facilitating diverse tasks and realizing the objectives outlined in the project blueprint. Encompassing both concrete elements such as hardware and intangible assets like software, resource requirements hold paramount importance. Effectively recognizing and overseeing these requisites assume a crucial role in guaranteeing the seamless progression of a project, the attainment of its set goals, and its successful culmination within the pre-established boundaries of scope and timeline.

Hardware:

Table 3.1.2.1 Hardware specification

Operating system	Windows 10 64-bit
Processor	Intel(R) Core (TM) i5-8265U CPU @ 1.60GHz 1.80 GHz
RAM	12 GB
Graphic card	NVIDIA GeForce MX250

The specified hardware configuration of a laptop, featuring 12 GB of RAM, an Intel i5-8265U CPU, and an NVIDIA GeForce MX250 graphics card enables the establishment of a solid groundwork for a diverse array of data analysis operations and computational undertakings. The 12 GB RAM allocation furnishes ample memory capacity which could effectively manage moderately sized datasets and intricate calculations with

notable efficiency. This aspect is particularly advantageous when engaged in association rule mining for market basket analysis where memory-intensive algorithms proficiently process expansive transactional datasets without encountering performance hindrances. The Intel i5-8265U CPU, hailing from Intel's 8th generation and possessing four cores achieves an equilibrium between performance prowess and energy efficiency. Boasting a base clock speed of approximately 1.6 GHz makes it capable of managing an assortment of computational responsibilities, encompassing data preprocessing, algorithmic execution, and visualization rendering. Meanwhile, although the NVIDIA GeForce MX250 graphics card is not classified as top-tier, however, it still furnishes substantial graphical processing potency for a laptop. This dedicated GPU significantly elevates the efficiency of data visualization tasks, facilitating smoother rendering of graphs, charts, and intricate visual depictions of association rules.

Software

1. Spyder

Spyder is an open-source integrated development environment (IDE) specifically designed for scientific computing, data analysis, and numerical computation. It is widely used by researchers and data analysts who work with Python for tasks such as machine learning and data exploration. One of the key features of Spyder is its ability to integrate with Python libraries used in data science and scientific computing. This integration allows users to easily import and manipulate data, visualize results, and perform complex computations within the same environment.

2. Seaborn, Matplotlib, and Pandas

The integration of libraries such as Seaborn, Matplotlib, and Pandas elevates the analytical process. Seaborn and Matplotlib facilitate data visualization by enabling the crafting of insightful plots and graphs that effectively illustrate association rules and trends in transactional data. Pandas is a robust data manipulation library that could be used to simplify the preprocessing of data using the functions provided. This streamlines the cleansing, transformation, and structuring of transactional data before it is channeled into the association rule mining algorithms.

3. Apriori and FP-Growth

In terms of algorithms, both Apriori and FP-Growth algorithms are used in the project. Apriori excels in unearthing association rules through the identification of frequent

itemsets whereas FP-Growth's forte lies in efficiently managing extensive datasets through a tree-based structure. These algorithms are harnessed within the Python environment, and analyze transactional data, eventually unveiling meaningful patterns, and interrelations between items. This analysis stands as the focal point of the project's core endeavors.

3.1.3 System performance definition

The definition of system performance within the project revolves around the precision and caliber of the association rules unearthed throughout the analysis. These rules must accurately portray meaningful patterns within customer purchasing behavior across distinct periods. Their significance, pertinence, and practical applicability stand as pivotal determinants of project success. Additionally, the efficacy of the Apriori and FP-Growth algorithms holds utmost importance. These algorithms drive the association rule mining process that might influence the project's capability to manage voluminous transactional datasets especially when dealing with different segmented time intervals. Equally vital is the efficiency aspect of system performance. This encompasses the computational efficiency of the algorithms to ensure the analysis's completion within reasonable timeframes. Given the potential complexity of item combinations to explore, the algorithms' speed profoundly impacts the project's feasibility. Furthermore, the judicious allocation of system resources, such as memory and processing prowess, fosters seamless execution and averts resource depletion during the analysis phase.

3.2 System Design / Overview

This subchapter will outline a detailed and methodical portrayal of the project's architecture, constituents, and interconnections will be presented. This presentation offers a high-level perspective of the amalgamation of diverse elements to attain the project's aims. The overview of the system design functions as a blueprint, elucidating the configuration of functional modules and the protocols for managing errors.

3.2.1 Functional modules

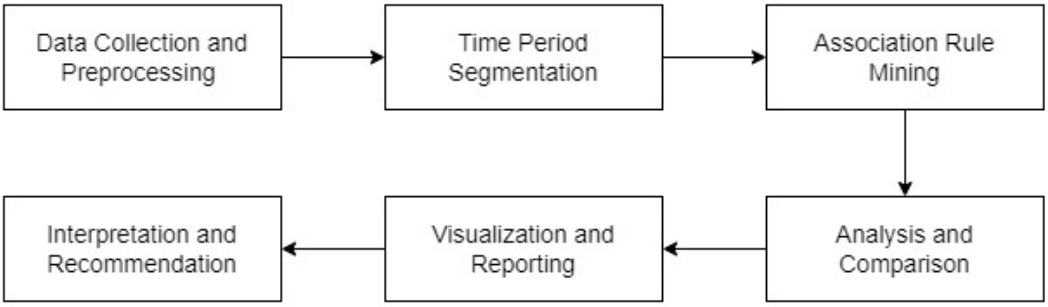


Figure 3.2.1.1. System module diagram

Primarily, the data collection and pre-processing module will concentrate on acquiring raw transactional data from pertinent sources such as retail databases, e-commerce platforms, or loyalty programs. This phase encompasses data extraction, refinement, transformation, and potentially integrating data from diverse origins. This module guarantees the precision, uniformity, and structured formatting of transactional data, rendering it suitable for subsequent analytical phases. The dataset to be used is a transactional dataset which contains all the transactions occurring between December of 2010 and 2011 and the source of the dataset originated from a UK-based and registered non-store online retail. The dataset contains a total of 541909 rows and 9 columns which correspond to invoice number, stock code, description of product, quantity, invoice date, customer code, country, and biweek.

Table 3.2.1.1. Structure of the transactional data

	InvoiceNo	StockCode	Descriptio	Quantity	InvoiceDate	UnitPrice	Customerl	Country	Biweek
2	536365	85123A	WHITE HA	6	12/1/2010 8:26	2.55	17850	United Kin	12/01/2010 - 12/14/2010
3	536365	71053	WHITE ME	6	12/1/2010 8:26	3.39	17850	United Kin	12/01/2010 - 12/14/2010
4	536365	84406B	CREAM CU	8	12/1/2010 8:26	2.75	17850	United Kin	12/01/2010 - 12/14/2010
5	536365	84029G	KNITTED U	6	12/1/2010 8:26	3.39	17850	United Kin	12/01/2010 - 12/14/2010
6	536365	84029E	RED WOO	6	12/1/2010 8:26	3.39	17850	United Kin	12/01/2010 - 12/14/2010
7	536365	22752	SET 7 BAB	2	12/1/2010 8:26	7.65	17850	United Kin	12/01/2010 - 12/14/2010
8	536365	21730	GLASS STA	6	12/1/2010 8:26	4.25	17850	United Kin	12/01/2010 - 12/14/2010
9	536366	22633	HAND WA	6	12/1/2010 8:28	1.85	17850	United Kin	12/01/2010 - 12/14/2010
10	536366	22632	HAND WA	6	12/1/2010 8:28	1.85	17850	United Kin	12/01/2010 - 12/14/2010
11	536367	84879	ASSORTED	32	12/1/2010 8:34	1.69	13047	United Kin	12/01/2010 - 12/14/2010
12	536367	22745	POPPY'S PI	6	12/1/2010 8:34	2.1	13047	United Kin	12/01/2010 - 12/14/2010
13	536367	22748	POPPY'S PI	6	12/1/2010 8:34	2.1	13047	United Kin	12/01/2010 - 12/14/2010
14	536367	22749	FELTCRAF	8	12/1/2010 8:34	3.75	13047	United Kin	12/01/2010 - 12/14/2010
15	536367	22310	IVORY KNI	6	12/1/2010 8:34	1.65	13047	United Kin	12/01/2010 - 12/14/2010
16	536367	84969	BOX OF 6 /	6	12/1/2010 8:34	4.25	13047	United Kin	12/01/2010 - 12/14/2010
17	536367	22623	BOX OF VI	3	12/1/2010 8:34	4.95	13047	United Kin	12/01/2010 - 12/14/2010
18	536367	22622	BOX OF VI	2	12/1/2010 8:34	9.95	13047	United Kin	12/01/2010 - 12/14/2010
19	536367	21754	HOME BUI	3	12/1/2010 8:34	5.95	13047	United Kin	12/01/2010 - 12/14/2010
20	536367	21755	LOVE BUIL	3	12/1/2010 8:34	5.95	13047	United Kin	12/01/2010 - 12/14/2010
21	536367	21777	RECIPE BO	4	12/1/2010 8:34	7.95	13047	United Kin	12/01/2010 - 12/14/2010
22	536367	48187	DOORMAT	4	12/1/2010 8:34	7.95	13047	United Kin	12/01/2010 - 12/14/2010
23	536368	22960	JAM MAKI	6	12/1/2010 8:34	4.25	13047	United Kin	12/01/2010 - 12/14/2010
24	536368	22913	RED COAT	3	12/1/2010 8:34	4.95	13047	United Kin	12/01/2010 - 12/14/2010
25	536368	22912	YELLOW C	3	12/1/2010 8:34	4.95	13047	United Kin	12/01/2010 - 12/14/2010
26	536368	22914	BLUE COA	3	12/1/2010 8:34	4.95	13047	United Kin	12/01/2010 - 12/14/2010
27	536369	21756	BATH BUIL	3	12/1/2010 8:35	5.95	13047	United Kin	12/01/2010 - 12/14/2010
28	536370	22728	ALARM CL	24	12/1/2010 8:45	3.75	12583	France	12/01/2010 - 12/14/2010
29	536370	22727	ALARM CL	24	12/1/2010 8:45	3.75	12583	France	12/01/2010 - 12/14/2010
30	536370	22726	ALARM CL	12	12/1/2010 8:45	3.75	12583	France	12/01/2010 - 12/14/2010
31	536370	21724	PANDA AN	12	12/1/2010 8:45	0.85	12583	France	12/01/2010 - 12/14/2010

For the pre-processing steps, the ‘InvoiceDate’ column is converted to a datetime data type to perform further date-based operations. A new column ‘Quarter’ in the DataFrame is created by extracting the quarter information from the ‘InvoiceDate’ column. The code then calculates the number of transactions for each quarter and year combination by grouping the data according to ‘Year’ and ‘Quarter’. Afterwards, the sum of transactions counts is calculated for each year’s four consecutive quarters and the top four quarters with the highest total transaction counts are selected. The code then filters the original Dataframe ‘df’ to retain only the transactions from the selected four quarters and eventually the filtered DataFrame ‘df_top_four_quarters’ is grouped by the ‘InvoiceNo’, ‘Year’, ‘Quarter’, and the ‘Description’ column is aggregated into a list for each group. This results in a new DataFrame ‘data_grouped’ where each row represents a unique transaction. Consequently, an empty list named ‘top_quarter_transactions’ is created to store the transactions. The code then iterates through the top four quarters which are sorted by their quarterly period and the ‘data_grouped’ DataFrame is filtered to extract only the transactions corresponding to that quarter for each quarter of the selected four quarters. The transactions for each quarter are assigned to separate variables for later processing operations purpose.

```
# Convert the "invoicedate" column to datetime data type
df['InvoiceDate'] = pd.to_datetime(df['InvoiceDate'])

# Categorize the data into different period data
df['Year'] = df['InvoiceDate'].dt.year

# Categorize into quarterly period
df['Quarter'] = df['InvoiceDate'].dt.to_period('Q')

# Select four quarters with the highest number of transactions
quarterly_counts = df.groupby(['Year', 'Quarter']).size().reset_index(name='Count')
quarterly_counts['ConsecutiveCount'] = (
    quarterly_counts.sort_values(['Year', 'Quarter'])
                    .groupby('Year')['Count']
                    .rolling(4, min_periods=1)
                    .sum()
                    .reset_index(drop=True)
)

top_four_quarters = quarterly_counts.sort_values('ConsecutiveCount', ascending=False).head(4)

# Filter the DataFrame to keep only the data from the top four quarters
df_top_four_quarters = df.merge(top_four_quarters[['Year', 'Quarter']], on=['Year', 'Quarter'])
data_grouped = df_top_four_quarters.groupby(['InvoiceNo', 'Year', 'Quarter'])['Description'].apply(list).reset_index()

top_quarter_transactions = []

# Iterate through the top four quarters and extract transactions
for top_quarter in sorted(top_four_quarters['Quarter']):
    top_quarter_df = data_grouped[data_grouped['Quarter'] == top_quarter]
    transactions = top_quarter_df['Description'].tolist()
    top_quarter_transactions.append(transactions)

quarter1 = top_quarter_transactions[0]
quarter2 = top_quarter_transactions[1]
quarter3 = top_quarter_transactions[2]
quarter4 = top_quarter_transactions[3]
```

Figure 3.2.1.2. Pre-processing of quarterly data

Similarly, three consecutive months with the highest number of transactions are selected for the monthly MBA to discover the most significant association rules. The DataFrame ‘df’ is firstly grouped by both 'Year' and 'Month' and counts the number of transactions in each month. The result is stored in the ‘monthly_counts’ DataFrame and the cumulative sum of transaction counts for every three consecutive months for each year is then calculated by using the ‘rolling’ method with a window size of 3. Eventually, the top three consecutive months with the highest cumulative transaction counts are identified and selected using the ‘ConsecutiveCount' column. The DataFrame ‘df’ is further filtered into another DataFrame, ‘top_three_months’ that stores information about the top three consecutive months with the highest transaction counts which include the ‘Year’ and ‘Month’ to retain only the rows that match the 'Year' and 'Month' combinations present in the recognized top three months. After filtering, the ‘df_top_three_months’ DataFrame is grouped by ‘InvoiceNo’, ‘Year’, and ‘Month’. The ‘list’ function is applied to the 'Description' column within each group for effectively aggregating all descriptions (items) associated with each invoice within the same month-year combination. The resulting DataFrame, ‘data_grouped’ now contains information about each invoice in the top three months and the 'Description' column stores a list of items in that invoice. The transactions for each of the top three months are assigned to separate variables for further analysis and insights into purchasing patterns during these periods.

```
# Select three consecutive months with the highest number of transactions
monthly_counts = df.groupby(['Year', 'Month']).size().reset_index(name='Count')
monthly_counts['ConsecutiveCount'] = (
    monthly_counts.sort_values(['Year', 'Month'])
                  .groupby('Year')['Count']
                  .rolling(3, min_periods=1)
                  .sum()
                  .reset_index(drop=True)
)

top_three_months = monthly_counts.sort_values('ConsecutiveCount', ascending=False).head(3)

# Filter the DataFrame to keep only the data from the top three months
df_top_three_months = df.merge(top_three_months[['Year', 'Month']], on=['Year', 'Month'])
data_grouped = df_top_three_months.groupby(['InvoiceNo', 'Year', 'Month'])['Description'].apply(list).reset_index()

top_month_transactions = []

# Iterate through the top three months and extract transactions
for top_month in sorted(top_three_months['Month']):
    top_month_df = data_grouped[data_grouped['Month'] == top_month]
    transactions = top_month_df['Description'].tolist()
    top_month_transactions.append(transactions)

month1 = top_month_transactions[0]
month2 = top_month_transactions[1]
month3 = top_month_transactions[2]
```

Figure 3.2.1.3. Pre-processing of monthly data

Regarding the pre-processing of biweekly data, the transactional data for biweekly MBA is selected from the top three consecutive months with the highest cumulative transaction counts which were identified previously. To realize the preparation of the biweekly data, the ‘df_top_three_months’ DataFrame is grouped by 'InvoiceNo' and 'Biweek', and the ‘list’ function is applied for each unique combination of 'InvoiceNo' and 'Biweek’ to the 'Description' column. The purpose of that operation is to aggregate all the descriptions (items) associated with each invoice within the same biweek. After applying the aggregation function, it resets the index of the resulting DataFrame, creating a new DataFrame named ‘biweekly_data_grouped’. The code then iterates through each biweek in ‘biweekly_data_grouped’ to calculate the transaction count for each biweek and stores the count in ‘biweek_transaction_counts’ which is initialized at the beginning. The transactions for each biweek also stored in ‘biweek_transaction’ and sorted in chronological order which is based on the start date of each biweek. Finally, the transactions for each of the six biweeks are assigned to separate variables.

```
# Initialize a dictionary to store biweek transaction counts
biweek_transaction_counts = defaultdict(int)
biweek_transactions = defaultdict(list)

biweekly_data_grouped = df_top_three_months.groupby(['InvoiceNo', 'Biweek'])['Description'].apply(list).reset_index()

# Sort biweekly_data_grouped by 'Biweek' column
biweekly_data_grouped = biweekly_data_grouped.sort_values('Biweek')

# Iterate through each biweek and calculate transaction counts
for biweek, group in biweekly_data_grouped.groupby('Biweek'):
    transaction_count = len(group)
    biweek_transaction_counts[biweek] = transaction_count
    transactions = group['Description'].tolist()
    biweek_transactions[biweek] = transactions

sorted_biweeks = sorted(biweek_transaction_counts.keys(), key=lambda biweek: pd.to_datetime(biweek.split(' - ')[0]))

# Iterate through the sorted order of biweeks
for biweek in sorted_biweeks:
    transactions = biweek_transactions[biweek]
    print(f"Number of transactions in {biweek}: {len(transactions)}")
    print()

biweek1 = biweek_transactions[sorted_biweeks[0]]
biweek2 = biweek_transactions[sorted_biweeks[1]]
biweek3 = biweek_transactions[sorted_biweeks[2]]
biweek4 = biweek_transactions[sorted_biweeks[3]]
biweek5 = biweek_transactions[sorted_biweeks[4]]
biweek6 = biweek_transactions[sorted_biweeks[5]]
```

Figure 3.2.1.4. Pre-processing of biweekly data

Additionally, another function ‘calculate_top_item_metrics’ is defined to compute various metrics for the top items in a given DataFrame. It will first calculate the count of each item in the ‘Description’ column and selects the top 20 items based on these counts. The function extracts a subset of the DataFrame containing only transactions involving the item in the top items list and calculates the total number of receipts, total sales quantity, and total sales value for that item. The support for the item

is computed as the ratio of total receipts for the item to the total number of invoices in the corresponding period which is retrieved from 'total_invoices_per_period'. The main objective of this function is to perform Exploratory Data Analysis (EDA) for further understanding of the most popular items in the specific period.

```
# Function to calculate metrics for the top items in a given DataFrame
def calculate_top_item_metrics(df, total_invoices_per_period, period_column, top_n=20):
    top_item_counts = df['Description'].value_counts().head(top_n)

    # Initialize lists to store metrics for each item
    top_items_metrics = []

    for item in top_item_counts.index:
        item_df = df[df['Description'] == item]
        total_receipts = len(item_df)
        total_sales_quantity = item_df['Quantity'].sum()
        total_sales_value = (item_df['Quantity'] * item_df['UnitPrice']).sum()

        # Calculate support for the item based on total invoices
        period = item_df[period_column].iloc[0]
        total_invoices = total_invoices_per_period[period]
        support = total_receipts / total_invoices

        top_items_metrics.append((item, total_receipts, total_sales_quantity, total_sales_value, support))

    return top_items_metrics
```

Figure 3.2.1.5. Calculation of metrics for top items

On top of the function for calculating metrics for the top items, an additional function 'process_and_visualize_top_items' is implemented for visualization purpose. This function takes three arguments which are the DataFrame containing transaction data, the DataFrame containing the total number of invoices in each period, and the name of the column in 'df' representing the period which could be either 'Quarter', 'Month', or 'Biweek'. It first processes the transaction data for each period and calculates metrics for the top items in each period using the 'calculate_top_item_metrics' function. The support values of the top items will be visualized in horizontal bar charts for users to gain better insights into the most popular items in that particular period.

```
def process_and_visualize_top_items(df, total_invoices_per_period, period_name):
    support_values = []

    # Process transactions and calculate metrics for each period
    for i, period in enumerate(total_invoices_per_period.index):
        period_df = df[df[period_name] == period]
        top_items_metrics = calculate_top_item_metrics(period_df, total_invoices_per_period, period_name)

        # Sort items by support value in descending order
        top_items_metrics.sort(key=lambda x: x[4], reverse=True)

        # Extract support values and item names
        item_names = [item for item, _, _, _, support in top_items_metrics]
        support_values_period = [support for _, _, _, _, support in top_items_metrics]
        support_values.append(support_values_period)

        # Visualization
        plt.figure(figsize=(10, 6))

        # Plotting in descending order
        plt.barh(item_names[::-1], support_values_period[::-1])

        plt.xlabel("Support")
        plt.ylabel("Item Name")
        plt.title(f"Support of Top 20 Items in {period_name.capitalize()} {i+1} ({period})")
        plt.tight_layout()
        plt.show()
```

Figure 3.2.1.6. Visualization of support of top items

To be mentioned that, there is a pre-requisite step for performing association rule mining, which is one-hot encoding on transaction data. One-hot encoding is a popular technique used in data pre-processing to convert categorical variables into a binary representation that can be used by machine learning algorithms. A function 'perform_one_hot_encoding' which takes a list of transactions and an optional parameter 'remove_lower' is defined. The function will remove columns containing any lowercase letters from the column names if 'remove_lower' is set to True as these typically represent noise or insignificant items in the transaction data. The function converts the transaction list into a list of strings and initializes a 'TransactionEncoder' object from the 'mlxtend.preprocessing' module which is used to perform one-hot encoding on the transaction data. The output will be each column representing an item and each row represents a transaction where value of 1 indicates the presence of an item in a transaction and value of 0 indicates absence. Later, a list 'transaction_sets' is defined to store the transactions for each of the quarterly, monthly, and biweekly periods. The encoding function is applied to each set of the transactions through a loop and store the resulting hot encoded DataFrames in a list 'one_hot_encoded_dfs'.

```
def perform_one_hot_encoding(transaction_list, remove_lower=False):
    transaction_list_str = [[str(item) for item in sublist] for sublist in transaction_list]

    encoder = TransactionEncoder()
    one_hot_encoded = encoder.fit_transform(transaction_list_str)
    one_hot_encoded_df = pd.DataFrame(one_hot_encoded, columns=encoder.columns_)

    if remove_lower:
        columns_to_remove = [col for col in one_hot_encoded_df.columns if any(c.islower() for c in col)]
        one_hot_encoded_df = one_hot_encoded_df.drop(columns=columns_to_remove)

    return one_hot_encoded_df

# Define the list of transaction sets
transaction_sets = [quarter1, quarter2, quarter3, quarter4, month1, month2, month3, biweek1, biweek2, biweek3, biweek4, biweek5, biweek6]

one_hot_encoded_dfs = []
for i, transactions in enumerate(transaction_sets, start=1):
    one_hot_encoded_dfs.append(perform_one_hot_encoding(transactions, remove_lower=True))
    globals()[f'one_hot_encoded_df{i}'] = one_hot_encoded_dfs[-1]
```

Figure 3.2.1.7. One-hot encoding of the data

At the core of the project, the association rule mining module employs algorithms such as Apriori and FP-Growth to extract association rules within each segmented time frame. This phase involves unearthing frequent itemsets and generating association rules based on predefined support and confidence thresholds. These processes unveil insights into item co-occurrence trends within transactions and uncover rules emblematic of purchasing behaviours specific to distinct periods. In the subsequent analysis and comparison module, the derived association rules from diverse time segments undergo assessment and juxtaposition. This module evaluates the weight, potency, and patterns of the unearthed rules. By contrasting rules across

disparate temporal divisions, trends, and deviations within customer purchasing conduct are revealed. This evaluation assumes paramount importance in pinpointing the optimal period for efficient association rule mining. The Apriori algorithm is a classic data mining technique used for association rule mining to discover interesting relationships between items in large datasets, particularly in Market Basket Analysis. The algorithm works based on the principle that if an itemset is frequent, then all of its subsets must also be frequent. The key metrics used in Apriori are support, confidence, and lift. Support measures the proportion of transactions in the dataset that contain a specific itemset by calculating the ratio of transactions containing the itemset to the total number of transactions. Confidence assesses the likelihood that an item Y is purchased when item X is bought. It is computed by discovering the ratio of the support of the itemset {X, Y} to the support of itemset X. Lastly, lift measures the strength of association between items X and Y by considering how much more likely Y is purchased when X is bought, compared to the case where they are bought independently. The equations for computing support, confidence, and lift are provided below:

$$s = \frac{\sum(Ta + Tb)}{\sum(T)}$$

where:

S = Support value

$\Sigma(Ta + Tb)$ = Transactions that contains item A (antecedent) and item B (consequent)

$\Sigma(T)$ = Total transactions

$$C = \frac{\sum(Ta + Tb)}{\sum(Ta)}$$

S = Confidence value

$\Sigma(Ta + Tb)$ = Transactions that contains item A (antecedent) and item B (consequent)

$\Sigma(T)$ = Total transactions that contains A

$$Lift = \frac{Confidence\ (A \rightarrow B)}{Suport\ (B)}$$

The code snippet below shows a function that generates association rules from a given one-hot encoded DataFrame is defined as ‘generate_association_rules’. The minimum support threshold is set to 0.02 and the minimum lift threshold is set to 1. The ‘apriori’ function will return a DataFrame of frequent itemsets where each row represents a set of items that occur together frequently in transactions. The ‘association_rules’ function then generates association rules from the frequent itemsets where each rule consists of an antecedent and a consequent along with metrics such a s support, confidence, and lift. The ‘min_threshold’ parameter is used to filter out rules that do not meet the minimum lift threshold specified. The function will raise a ‘ValueError’ if no frequent itemsets are found to indicate that no rules could be generated. Otherwise, the function will return the DataFrame of association rules which will be used for further analysis or making recommendations in a market basket analysis context.

```
min_support = 0.02
min_threshold = 1

def generate_association_rules(one_hot_encoded_df, min_support=0.02, min_threshold=1):
    frequent_itemsets = apriori(one_hot_encoded_df, min_support=min_support, use_colnames=True)

    if frequent_itemsets.empty:
        raise ValueError("Frequent itemsets DataFrame is empty.")

    rules = association_rules(frequent_itemsets, metric="lift", min_threshold=min_threshold)
    return rules
```

Figure 3.2.1.8. Generation of association rules

Subsequently, the visualization and reporting module transforms the findings into comprehensible visual representations and reports. The generated reports provide a concise overview of the analysis results to facilitate decision-makers in gleaning insights and actionable recommendations. Subsequent to the analysis, the last module will interpret the ramifications of association rules uncovered within distinct time frames. This translation of patterns into meaningful insights establishes a foundation for formulating recommendations. The module identifies potential strategies for refining marketing endeavours, amplifying customer engagement, or honing inventory management strategies. The code is then responsible for exporting the results of association rule mining to an Excel file. It iterates through the association rules generated for the top three months and prints out the top 20 association rules along with additional information such as support, confidence, lift, total receipts, total sales quantity, and total sales value. The analysis will be carried out by focusing on the

changes in the metrics such as lift and support. Any significant fluctuation of the values will be observed and meaningful insights are extracted.

3.2.2 Error handling

In the domain of association rule mining, the minimum support threshold holds essential significance as it governs the frequency criterion that a set of items must fulfil to qualify as a frequent item set. Opting for a lower minimum support threshold can give rise to an abundance of candidate itemsets which will yield a substantial volume of association rules. However, this strategy can introduce several challenges such as notably memory space errors. To adeptly address these errors, meticulous adjustment of the minimum support threshold is essential. By slightly elevating the threshold, the proliferation of generated rules can be managed, hence, reducing the likelihood of memory space errors. This course of action ensures an equilibrium between capturing substantial insights and upholding the stability of the system.

3.3 Implementation Issue and Challenges

Issues arising during project implementation can delay the progress, compromise result integrity, and necessitate meticulous attention and resolution to ensure project fruition. One of the most substantial issues in this project is the minimum support threshold selection during the generation of association rules. Setting it too low can result in an excessive number of rules which makes it challenging to find meaningful insights. Conversely, setting it too high may lead to the exclusion of valuable rules. Hence, it took much time to find the right balance between comprehensiveness and relevance. Moreover, the preparation of the monthly data delays the overall progress of the project. This could be explained by the top three consecutive months with the highest cumulative transaction counts that might span across different years which might cause inaccurate extraction of data. The condition of allowing the months to span across years has to be specified in the coding work to ensure the accuracy of the data.

3.4 Timeline

3.4.1 FYP 2 Timeline

Table 3.4.1.1. FYP 2 Timeline

	Task	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13
1	Meet up with supervisor regarding FYP 2													
2	System design reevaluation													
3	System testing													
4	System debugging													
5	Report writing													
6	Report submission													
7	Presentation													

CHAPTER 4 EXPERIMENTAL RESULTS AND DISCUSSION

In this chapter, a detailed analysis of the market basket analysis results across quarterly, monthly, and biweekly periods will be presented. The discussion will include the top 20 highest frequency itemsets and product affinity analysis for each period. This analysis offers valuable insights into customer behavior and purchasing patterns. The discussion of the top 20 highest frequency itemsets in each period would reveal the most popular products and their combinations which allows businesses to obtain valuable information for optimizing product placement and marketing strategies. Additionally, the product affinity analysis will provide insights into the relationships between different products which may offer cross-selling opportunities. Last but not least, a comparison of MBA results in different periods will be carried out to highlight the advantages and disadvantages of each approach. The comparison will identify the optimal period for conducting an MBA by considering factors such as the frequency of data updates, the granularity of insights required, and the resources needed for analysis.

4.1 Quarterly Result

The quarterly market basket analysis is conducted using the quarterly transaction data of 2011 and the main objective of conducting market basket analysis for the quarterly period is to figure out if there was seasonality for sales of certain products. The first quarter of 2011 witnessed 4852 transactions and a transaction number of around 5000 reflecting a post-holiday period where customer spending tends to be more conservative after the previous peak shopping season year-end holidays especially Christmas. This trend could be explained with customers were trying to recover from holiday expenditures. Businesses are advised to leverage this insight to adjust their marketing strategies and promotions such as promoting early-bird deals for upcoming seasonal events or offering discounts on non-seasonal items. Afterwards, a significant surge in the number of transactions in the second quarter could be observed where the transactions increased to 5918. The around 20% increase in transactions indicates a resurgence in customer activity as the weather warms up and various springtime events and holidays approach. This uptick in spending could be attributed to people starting to participate in more outdoor activities, spring cleaning, and preparations for summer vacation. Businesses could capitalize on this trend by launching seasonal promotions for certain targeted products that resonate with the changing customer needs and preferences during this period. Moving

into the third quarter, the number of transactions further increased to 5991 which signals the continued growth in customer activity. This period conventionally included back-to-school shopping and end-of-summer clearance sales. Hence, back-to-school discounts for certain school-related products could be launched, and prepare new product lines for the upcoming fall season to meet the growing demand. Last but not least, the last quarter of 2011 experienced the highest number of transactions for the year with 7114 transactions. This surge in customer activity is reasonable due to this period covers certain major shopping events such as Black Friday, Cyber Monday, and Christmas. Therefore, businesses are advised to ensure sufficient inventory levels to meet the heightened demand and maximize the profit.

Table 4.1.1. Top 20 highest frequency itemsets in 2011 Quarter 1

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	REGENCY CAKESTAND 3 TIER	0.1103	535	3114	39039.14
2	WHITE HANGING HEART T-LIGHT HOLDER	0.1094	531	9386	25928.01
3	SET OF 3 CAKE TINS PANTRY DESIGN	0.1043	506	2584	12706.59
4	JUMBO BAG RED RETROSPOT	0.0868	412	10998	20598.92
5	SET OF 6 SPICE TINS PANTRY DESIGN	0.0761	369	1906	8034.34
6	JAM MAKING SET WITH JARS	0.0705	342	2352	9833.08
7	NATURAL SLATE HEART CHALKBOARD	0.0688	334	2166	6708.22
8	LUNCH BAG RED RETROSPOT	0.0666	323	3938	6828.94
9	HEART OF WICKER SMALL	0.0653	317	5170	9344.22
10	PACK OF 72 RETROSPOT CAKE CASES	0.0645	313	9479	5145.31
11	JAM MAKING SET PRINTED	0.0631	306	4348	6641.72
12	PARTY BUNTING	0.0631	306	3072	13734.53
13	RECIPE BOX PANTRY YELLOW DESIGN	0.0577	280	1522	4573.77
14	ASSORTED COLOUR BIRD ORNAMENT	0.0548	266	6087	9937.67
15	SET/20 RED RETROSPOT PAPER NAPKINS	0.0542	263	4164	3430.18
16	GREEN REGENCY TEACUP AND SAUCER	0.0538	261	1427	4616.30
17	SET OF 3 HEART COOKIE CUTTERS	0.0534	259	2110	2729.10
18	WOODEN FRAME ANTIQUE WHITE	0.0528	256	1632	5115.47
19	PLEASE ONE PERSON METAL SIGN	0.0523	254	2703	5459.73
20	VICTORIAN GLASS HANGING T-LIGHT	0.0519	252	5313	7023.46

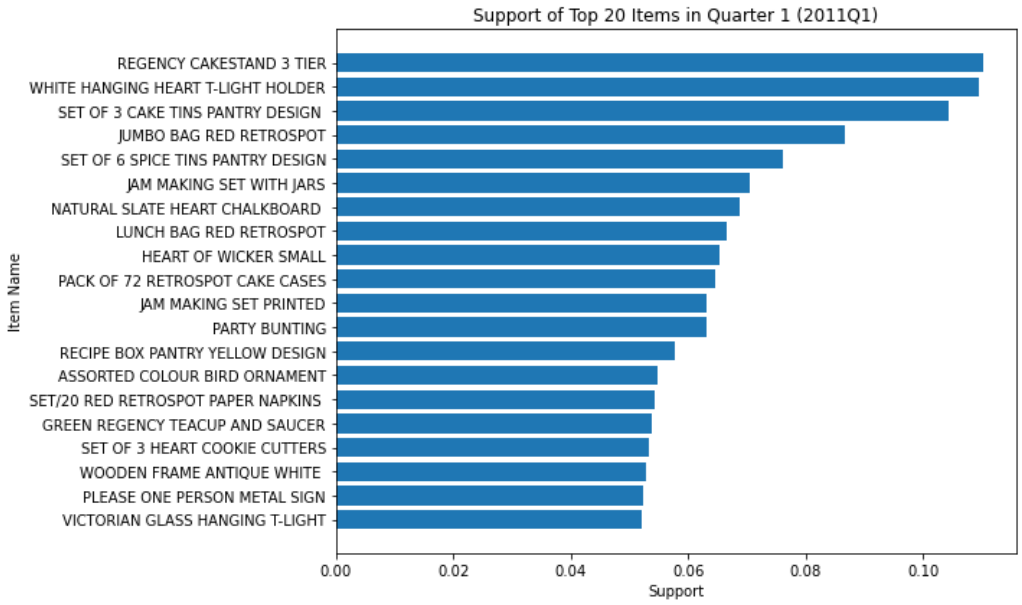


Figure 4.1.1. Visualization of support for top 20 popular items in Quarter 1

There are a few notable items that provide precious insights for businesses to understand customer preferences and purchasing trends during the first quarter of the year. First and foremost, the “REGENCY CAKESTAND 3 TIER” had the highest support of 0.1103 and took the top position in the ranking. Additionally, it had a total sales quantity of 3114 and a total sales value of 39039.14 which reflects this item was frequently purchased relative to other items in the dataset. Businesses are advised to promote this item to capitalize on its high demand and ensure sufficient stock levels for it. Another popular item is the “WHITE HANGING HEAR T-LIGHT HOLDER” which had a support of 0.1094. Although its support is slightly lower than the “REGENCY CAKESTAND 3 TIER”, however, it had a higher total sales quantity of 9386 which indicates a strong customer interest. The popularity of this item suggests a strong interest in decorative and ambient lighting products. Businesses can consider expanding their range of decorative lighting options to meet the demand for stylish home decoration items. The “SET OF 3 CAKE TINS PANTRY DESIGN” also performed well in the first quarter of the year. This item’s total sales value of 12706.59 indicates a strong market for kitchen storage solutions which suggests businesses enhance their offerings in this category by introducing more innovative and stylish storage solutions to attract more customers looking to organize their kitchens effectively.

Table 4.1.2. Product affinity analysis for 2011 Q1

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0221	0.9386	18.0003	675	4198	12867.73
2	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0221	0.4229	18.0003	675	4198	12867.73
3	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	0.0221	0.6564	17.6948	675	4198	12867.73
4	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0221	0.5944	17.6948	675	4198	12867.73
5	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SACUER	ROSES REGENCY TEACUP AND SAUCER	0.0221	0.8492	16.8178	675	4198	12867.73
6	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0221	0.4367	16.8178	675	4198	12867.73
7	ALARM CLOCK BAKELIKE RED	ALARM CLOCK BAKELIKE GREEN	0.0247	0.6283	16.2148	383	1665	6855.83
8	ALARM CLOCK BAKELIKE GREEN	ALARM CLOCK BAKELIKE RED	0.0247	0.6383	16.2148	383	1665	6855.83
9	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0260	0.7730	14.8246	425	2397	7562.17
10	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0260	0.4980	14.8246	425	2397	7562.17
11	DOLLY GIRL LUNCH BOX	SPACEBOY LUNCH BOX	0.0214	0.5714	14.4405	377	8079	14335.72
12	SPACEBOY LUNCH BOX	DOLLY GIRL LUNCH BOX	0.0214	0.5417	14.4405	377	8079	14335.72
13	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0371	0.7347	14.0899	511	3228	9921.86
14	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0371	0.7115	14.0899	511	3228	9921.86
15	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0235	0.6994	13.8507	414	2771	8251.43
16	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0235	0.4653	13.8507	414	2771	8251.43
17	RED RETROSPOT CHARLOTTE BAG	WOODLAND CHARLOTTE BAG	0.0214	0.5361	13.6182	390	7046	8585.76
18	WOODLAND CHARLOTTE BAG	RED RETROSPOT CHARLOTTE BAG	0.0214	0.5445	13.6182	390	7046	8585.76
19	JUMBO BAG WOODLAND ANIMALS	JUMBO BAG OWLS	0.0210	0.5455	12.7852	395	5228	10531.09
20	JUMBO BAG OWLS	JUMBO BAG WOODLAND ANIMALS	0.0210	0.4928	12.7852	395	5228	10531.09

According to the product affinity analysis conducted for the first quarter of the year, it highlights several interesting insights into customer behavior and product relationships and one of the notable patterns is the strong affinity between different variants of the “REGENCY TEACUP AND SAUCER” sets. For instance, one of the rules shows that customers who purchase the “ROSES REGENCY TEACUP AND SAUCER” along with the “PINK REGENCY TEACUP AND SAUCER” are highly likely to also buy the “GREEN REGENCY TEACUP AND SAUCER”. This association is supported by a high lift value of 18.0003 which indicates a significant relationship between these items and made this rule the top spot in the ranking of the first quarter. Given the strong association between different variants of the teacup sets, businesses are advised to consider bundling these items together as a set or offering discounts for purchasing multiple variants together. Moreover, another interesting insight is the association between “DOLLY GIRL LUNCH BOX” and “SPACEBOY LUNCH BOX”. The association rule between these items contributed a total sales value of 13335.72 in the first quarter and it was the highest sales value among the items in the ranking. Although this rule did not have a high position in the ranking, its incredible amount of total sales quantity and total sales value made it outperform other rules in this period. The association between the lunch boxes indicates these items appeal to a similar customer segment and businesses could take this opportunity to create themed sets or promotions that include both lunch boxes to boost sales. Similarly, there was an interesting cross-selling opportunity highlighted by the rule which included the items “BAKELIKE ALARM CLOCK RED” and “BAKELIKE ALARM CLOCK GREEN”. The association between these items indicated they might be complementary products that attract customers to purchase them for collection purposes. Businesses could create visually appealing displays in-store that showcase both colors of the “BAKELIKE ALARM CLOCK” and use eye-catching signage to draw attention to the cross-selling opportunity. Moreover, there was an association between the Charlotte bags in the first quarter as well. This is verified by the rule which contained “RED RETROSPOT CHARLOTTE BAG” and “WOODLAND CHARLOTTE BAG” with a high support value of 0.0214 and a lift of 13.6182. Businesses could introduce complementary accessories such as scarves or wallets along with the Charlotte bags to encourage customers to drive sales among fashion-forward customers.

Table 4.1.3. Top 20 highest frequency itemsets in 2011 Quarter 2

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PART BUNTING	0.1114	659	7589	41903.99
2	WHITE HANGING HEART T-LIGHT HOLDER	0.0960	568	7589	21100..94
3	REGENCY CAKESTAND 3 TIER	0.0918	543	2936	35964.56
4	JUMBO BAG RED RETROSPOT	0.0850	503	9745	19222.89
5	LUNCH BAG RED RETROSPOT	0.0705	417	5102	9661.86
6	SPOTTY BUNTING	0.0691	409	2764	13845.40
7	LUNCH BAG APPLE DESIGN	0.0673	398	5040	8060.39
8	ASSORTED COLOUR BIRD ORNAMENT	0.0600	355	7159	11811.75
9	JUMBO BAG APPLES	0.0588	348	5024	9870.12
10	PACK OF 72 RETROSPOT CAKE CASES	0.0586	347	8301	4886.52
11	LUNCH BAG SUKI DESIGN	0.0580	343	3307	5869.10
12	JUMBO STORAGE BAG SUKI	0.0563	333	2897	6806.81
13	JUMBO BAG PINK POLKADOT	0.0554	328	4123	8426.83
14	SET OF 4 PANTRY JELLY MOULDS	0.0544	322	3154	3946.83
15	JUMBO BAG DOILEY PATTERNS	0.0534	316	4435	8703.05
16	ROSES REGENCY TEACUP AND SAUCER	0.0532	315	2848	8121.60
17	PAPER BUNTING RETROSPOT	0.0526	311	3288	9959.28
18	LUNCH BAG DOILEY PATTERN	0.0517	306	3955	6323.91
19	JUMBO SHOPPER VINTAGE RED PAISLEY	0.0517	306	2951	6848.64
20	WOODEN PICTURE FRAME WHITE FINISH	0.0515	305	2105	5297.12

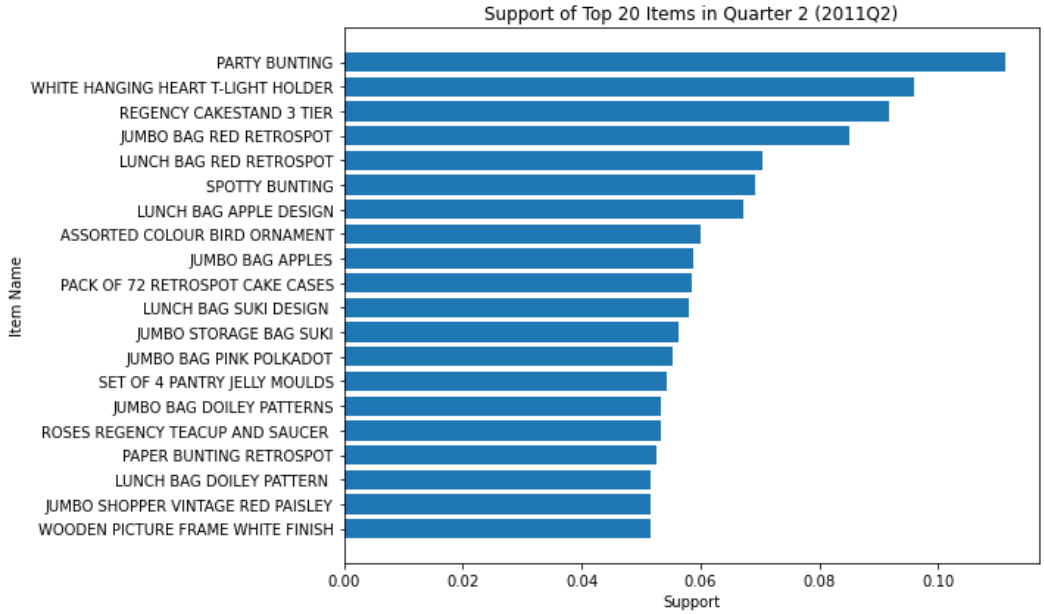


Figure 4.1.2. Visualization of support for top 20 popular items in Quarter 2

Comparing the top 20 popular items between the first quarter and the second quarter reveals certain significant insights and fluctuations in terms of sales metrics could be observed. One item that exhibited a substantial change in performance is “REGENCY CAKESTAND 3 TIER” which was the top-performing item in the previous quarter. Its support decreased from 0.1103 to 0.0918 which represents a decline of approximately 16.7%. Despite this, this item accumulated a total sales value of 35964.56 in the second quarter which suggests businesses could continue to promote this item as a stylish and practical accessory for hosting events and parties. Another notable change is the rise of “PARTY BUNTING” to the top spot in the second quarter with a support value of 0.1114. This item is witnessed with a substantial increase in total sales quantity to 7589 and total sales value to 41903.99 which is the highest sales volume in the ranking. According to this insight, businesses could consider expanding their range of party supplies and decorations to cater to the growing demand for event planning and celebration products. Additionally, “WHITE HANGING HEART T-LIGHT HOLDER” experienced a slight decrease in support from 0.1094 to 0.0960 in the second quarter while maintaining its top position in the ranking. The consistent performance of this item reflects a stable demand for decorative lighting products and businesses can continue to offer a variety of t-light holders and decorative lighting options to meet the stable demand for the items.

Table 4.1.4. Product affinity analysis for 2011 Q2

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	REGENCY TEA PLATE GREEN	REGENCY TEA PLATE ROSES	0.0206	0.8841	30.7756	311	6023	9457.52
2	REGENCY TEA PLATE ROSES	REGENCY TEA PLATE GREEN	0.0206	0.7176	30.7756	311	6023	9457.52
3	PINK VINTAGE PAISLEY PICNIC BAG	SCANDINAVIAN PAISLEY PICNIC BAG	0.0218	0.7207	23.9603	262	2899	5069.60
4	SCANDINAVIAN PAISLEY PICNIC BAG	PINK VINTAGE PAISLEY PICNIC BAG	0.0218	0.7247	23.9603	262	2899	5069.60
5	GARDENERS KNEELING PAD CUP OF TEA	GARDENERS KNEELING PAD KEEP CALM	0.0216	0.7151	21.2657	380	3182	5076.22
6	GARDENERS KNEELING PAD KEEP CALM	GARDENERS KNEELING PAD CUP OF TEA	0.0216	0.6432	21.2657	380	3182	5076.22
7	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0292	0.8122	20.5412	867	6211	18286.89
8	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	0.0292	0.7393	20.5412	867	6211	18286.89
9	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0292	0.6133	18.7446	867	6211	18286.89
10	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0292	0.8964	18.7446	867	6211	18286.89
11	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0330	0.8333	17.4264	552	3363	10165.29
12	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0330	0.6890	17.4264	552	3363	10165.29
13	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0292	0.8872	16.9914	867	6211	18286.89
14	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0292	0.5599	16.9914	867	6211	18286.89
15	STRAWBERRY CHARLOTTE BAG	RED RETROSPOT CHARLOTTE BAG	0.0206	0.6489	16.6251	429	6275	6837.98
16	RED RETROSPOT CHARLOTTE BAG	STRAWBERRY CHARLOTTE BAG	0.0206	0.5281	16.6251	429	6275	6837.98
17	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0326	0.8248	15.7964	570	4392	12744.56
18	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0326	0.6246	15.7964	570	4392	12744.56
19	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0360	0.7527	14.4148	612	4667	13663.93
20	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0360	0.6893	14.4148	612	4667	13663.93

In the second quarter, the first and second place of the ranking were taken by the itemsets "REGENCY TEA PLATE GREEN" and "REGENCY TEA PLATE ROSES" which exhibits a strong association between these two items with a high lift value of 30.7756 and a support value of 0.0206. Similarly, businesses could create bundled offers or promotions that encourage customers to purchase both variants together and highlight the complementary nature of the products to boost the average transaction value and enhance customer satisfaction. The previous dominant of the top spot in the ranking which was the rule involving "ROSES REGENCY TEACUP AND SAUCER", "PINK REGENCY TEACUP AND SAUCER", and "GREEN REGENCY TEACUP AND SAUCER" dropped to tenth due to the entries of stronger competitors in the ranking. The different variants of the teacup sets had lifts ranging from 18.7446 to 30.7756 which indicates that they were still attracting customers to purchase them together. Businesses could identify customers who have purchased one variant and offer them personalized recommendations or discounts on the other variants to encourage repeat purchases and increase customer loyalty. Another noteworthy insight is the association between the "PINK VINTAGE SCANDINAVIAN PAISLEY PICNIC BAG" and "SCANDINAVIAN PAISLEY PICNIC BAG" with a lift of 23.9603. The association indicating customers who purchase one variant are highly likely to purchase the other. Businesses could leverage this affinity to create themed picnic sets that include both variants along with other picnic accessories and target outdoor enthusiasts and picnic lovers to enhance the overall customer experience and drive sales. The analysis also highlights the affinity between different variants of the "GARDENERS KNEELING PADS" such as the "GARDENERS KNEELING PAD CUP OF TEA" and "GARDENERS KNEELING PAD KEEP CALM" with a lift of 21.2657. Bundling these two variants together as part of a gardening essentials kit and offering them at a special price could encourage customers to purchase both and increase the perceived value of the offer. Businesses could also design theme-based marketing campaigns centered around the kneeling pad variants to highlight the unique features and benefits of each variant or the practicality for different gardening tasks. This is to tailor marketing messages to resonate with specific customer segments such as gardening enthusiasts or outdoor hobbyists.

Table 4.1.5. Top 20 highest frequency itemsets in 2011 Quarter 3

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	JUMBO BAG RED RETROSPOT	0.0995	596	12794	25498.01
2	WHITE HANGING HEART T-LIGHT HOLDER	0.0868	520	7584	21370.17
3	SPOTTY BUNTING	0.0833	499	4114	20204.50
4	PARTY BUNTING	0.0821	492	5519	30365.25
5	LUNCH BAG RED RETROSPOT	0.0795	476	5396	9573.11
6	REGENCY CAKESTAND 3 TIER	0.0774	464	2532	31498.89
7	SET OF 3 REGENCY CAKE TINS	0.0736	441	2816	13738.80
8	LUNCH BAG APPLE DESIGN	0.0683	409	4517	7950.87
9	LUNCH BAG BLACK SKULL	0.0683	409	4246	7843.13
10	LUNCH BAG SUKI DESIGN	0.0681	408	4239	7667.40
11	JUMBO BAG APPLES	0.0673	403	5979	12263.21
12	LUNCH BAG CARS BLUE	0.0634	380	4147	7415.17
13	LUNCH BAG PINK POLKADOT	0.0629	377	3649	6068.28
14	JUMBO BAG VINTAGE LEAF	0.0618	370	4595	9142.05
15	SET OF 3 CAKE TINS PANTRY DESIGN	0.0618	370	2344	11872.76
16	ASSORTED COLOUR BIRD ORNAMENT	0.0606	363	11444	17974.84
17	JUMBO BAG PINK POLKADOT	0.0594	356	6521	13237.93
18	JUMBO BAG ALPHABET	0.0586	351	5122	10119.86
19	LUNCH BAG SPACEBOY DESIGN	0.0578	346	3284	5732.48
20	JUMBO STORAGE BAG SUKI	0.0564	338	4001	9090.18

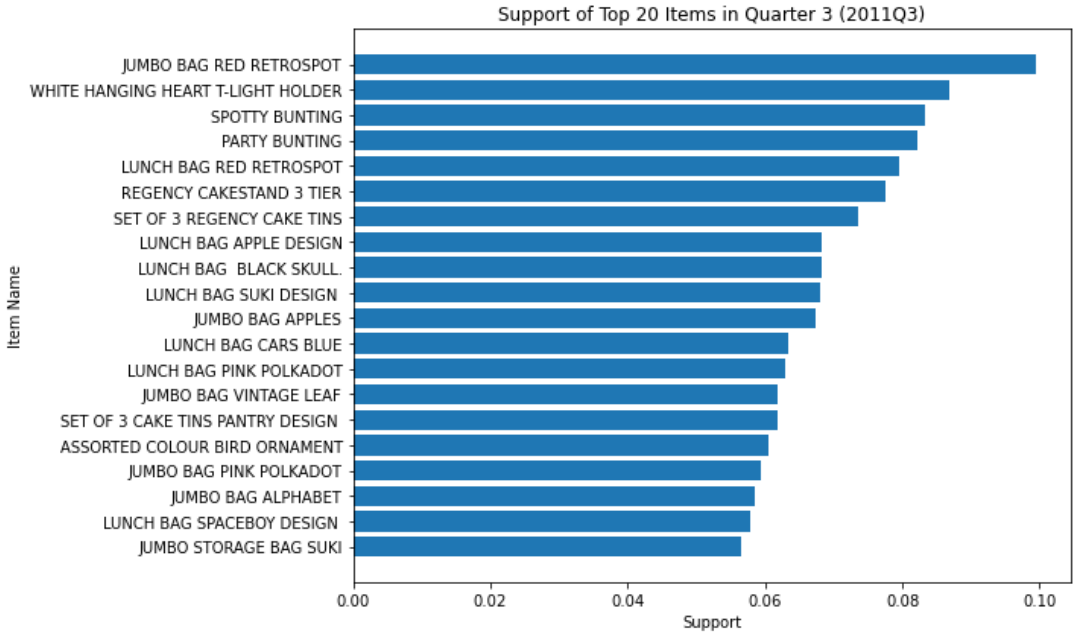


Figure 4.1.3. Visualization of support for top 20 popular items in Quarter 3

According to the top 20 items in the third quarter of the year, one notable change is the decrease in support for the "PARTY BUNTING" which was the top item in the previous quarter but it dropped to the fourth position in the third quarter with support of 0.0821 which represents a decline of approximately 26.2%. Despite the support of this item decreased, however, the total sales quantity and total sales value remained relatively stable which indicates continued popularity among customers. On the other hand, the "JUMBO RED RETROSPOT" experienced a rise in support and made it to the top spot in the ranking of the third quarter. The support of this item increased by approximately 17.7% as well as the total sales quantity increased by 31.3% compared to the previous quarter, reaching 12794 units while the total sales value increased by 32.6% to 25498.01. This indicates a growing demand for the "JUMBO BAG RED RETROSPOT" and businesses may consider increasing their inventory of this item to meet the rising demand. Another item that saw a significant increase in both support and ranking is "SPOTTY BUNTING" which rose from the sixth position in the previous quarter to the third position in the third quarter. The rise in ranking is contributed by the 32.2% increase in support and the total sales quantity of this item also increased by 49.2% to 4114 units while the total sales value increased by 46.0% to 20204.60. Businesses could promote this item more prominently to capitalize on its growing demand.

Table 4.1.6. Product affinity analysis for 2011 Q3

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0234	0.5738	21.6192	733	6523	19196.61
2	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0234	0.8805	21.6192	733	6523	19196.61
3	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0234	0.7735	21.5531	733	6523	19196.61
4	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	0.0234	0.6512	21.5531	733	6523	19196.61
5	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0285	0.7953	19.5284	463	3839	11401.45
6	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0285	0.7008	19.5284	463	3839	11401.45
7	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SACUER	0.0234	0.5224	18.3019	733	6523	19196.61
8	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SACUER	ROSES REGENCY TEACUP AND SAUCER	0.0234	0.8187	18.3019	733	6523	19196.61
9	DOLLY GIRL LUNCH BOX	SPACEBOY LUNCH BOX	0.0239	0.6530	17.7815	448	6374	11615.95
10	SPACEBOY LUNCH BOX	DOLLY GIRL LUNCH BOX	0.0239	0.6500	17.7815	448	6374	11615.95
12	SET OF 6 SNACK LOAF BAKING CASES	SET OF 6 TEA TIME BAKING CASES	0.0204	0.6321	17.2139	417	6256	6189.17
12	SET OF 6 TEA TIME BAKING CASES	SET OF 6 SNACK LOAF BAKING CASES	0.0204	0.5525	17.2139	417	6256	6189.17
13	STRAWBERRY CHARLOTTE BAG	WOODLAND CHARLOTTE BAG	0.0207	0.6078	16.7045	427	6099	7021.70
14	WOODLAND CHARLOTTE BAG	STRAWBERRY CHARLOTTE BAG	0.0207	0.5688	16.7045	427	6099	7021.70
15	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0302	0.7418	16.5826	515	4671	13760.93
16	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0302	0.6754	16.5826	515	4671	13760.93
17	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0265	0.7395	16.5319	488	4536	13230.84
18	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0265	0.5933	16.5319	488	4536	13230.84
19	SET OF 6 SNACK LOAF BAKING CASES	SET OF 12 MINI LOAF BAKING CASES	0.0214	0.6632	16.4867	441	7031	5645.50
20	SET OF 12 MINI LOAF BAKING CASES	SET OF 6 SNACK LOAF BAKING CASES	0.0214	0.5311	16.4867	441	7031	5645.50

Based on the product affinity analysis for the third quarter of the year, there were shifts in ranking for certain rules and one of the most obvious observations would be the top spot of the ranking was taken by the rule with the itemsets “GREEN REGENCY TEACUP AND SAUCER”, “ROSES REGENCY TEACUP AND SAUCER”, and “PINK REGENCY TEACUP AND SAUCER”. This rule obtained the highest lift among the rules in the ranking with a lift value of 21.6192 while having a declined support of 0.0234. The rule involving “REGENCY TEA PLATE GREEN” and “REGENCY TEA PLATE ROSES” which was the first place in the ranking of the previous quarter was eliminated by other rules with higher lift. There were over half rules in the ranking related to the teacup and saucer sets while the only difference between them was the colour variants. Aside from teacup and saucer sets, there were other rules related to Charlotte bags and baking cases as well. For instance, there was an association between “ SET OF 6 SNACK LOAF BAKING CASES” and “SET OF 6 TEA TIME BAKING CASES” with a relatively high lift of 17.2139 and a support of 0.0204. For the Charlotte bags, the item “STRAWBERRY CHARLOTTE BAG” was no longer being purchased with “RED RETROSPOT CHARLOTTE BAG” but “WOODLAND CHARLOTTE BAG”. The new association between the Charlotte bags showed up in the ranking with a slightly higher lift of 16.7045 compared to the previous rule with a different consequent which had a lift of 16.6251. Businesses could leverage the association between the Charlotte bags and other related products to engage with customers through social media and email marketing to promote the versatility and style of the Charlotte bags. They might also consider introducing new Charlotte bags in different settings and styles to reach a wider audience and drive sales. Another notable observation is the return of the association rule involving “DOLLY GIRL LUCNH BOX” and “SPACEBOY LUNCH BOX” with a lift of 17.7815 and support of 0.0239. Based on this insight, businesses should implement effective inventory management practices to ensure the availability of lunch boxes throughout the year and consider using demand forecasting techniques to predict future sales and adjust stock levels accordingly. They should also ensure the promotional activities for lunch boxes are consistent across quarters to maintain customer interest and encourage repeat purchases.

Table 4.1.7. Top 20 highest frequency itemsets in 2011 Quarter 4

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PAPER CHAIN KIT 50'S CHRISTMAS	0.1075	765	13146	45173.14
2	RABBIT NIGHT LIGHT	0.1036	737	25500	56289.81
3	HOT WATER BOTTLE KEEP CALM	0.0828	589	4280	21542.24
4	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0752	535	6293	20226.77
5	JUMBO BAG RED RETROSPOT	0.0723	514	11700	23063.86
6	JUMBO BAG 50'S CHRISTMAS	0.0720	512	7319	17757.84
7	WHITE HANGING HEART T-LIGHT HOLDER	0.0715	509	7415	21391.53
8	POPCORN HOLDER	0.0694	494	24222	24157.37
9	REGENCY CAKESTAND 3 TIER	0.0656	467	2436	31362.24
10	HAND WARMER OWL DESIGN	0.0641	456	4969	10795.34
11	GARDENERS KNEELING PAD KEEP CALM	0.0628	447	4492	9341.16
12	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0616	438	8949	3159.77
13	JUMBO BAG VINTAGE DOLLY	0.0595	423	6314	14352.96
14	CHOCOLATE HOT WATTER BOTTLE	0.0593	422	2904	15410.43
15	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0589	419	8746	3082.80
16	BAKING SET 9 PIECE RETROSPOT	0.0586	417	2035	10124.97
17	HOT WATER BOTTLE TEA AND SYMPATHY	0.0586	417	3259	18798.95
18	ASSORTED COLOUR BIRD ORNAMENT	0.0579	412	9320	15249.60
19	SET OF 20 VINTAGE CHRISTMAS NAPKINS	0.0559	398	4714	4557.08
20	VINTAGE DOILY TRAVEL SEWING KIT	0.0555	395	3695	7960.33

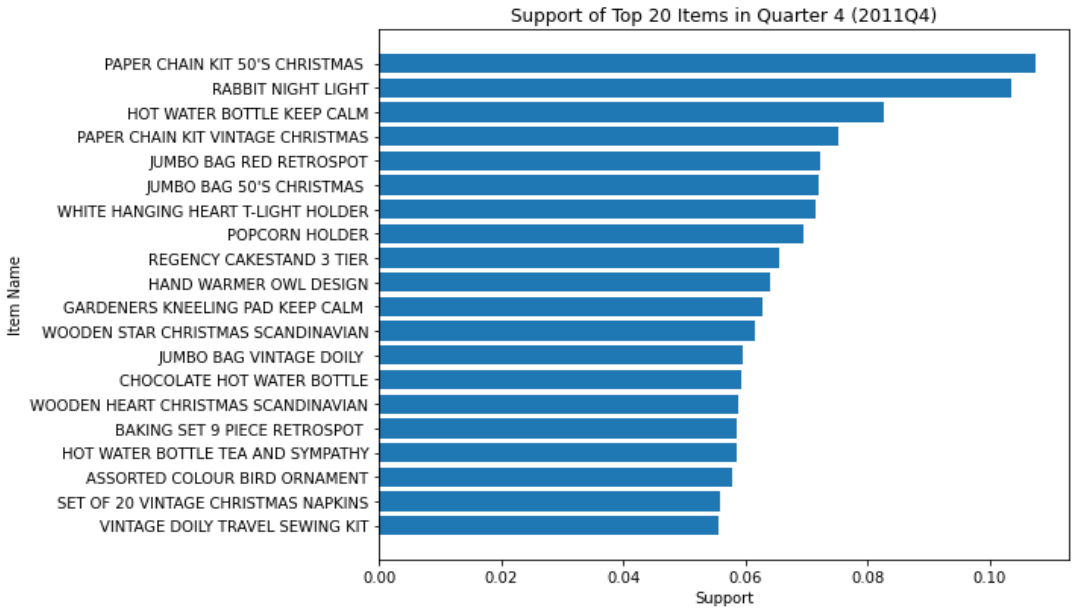


Figure 4.1.4. Visualization of support for top 20 popular items in Quarter 4

As stated in the table above that shows the top 20 items in the fourth quarter of the year, there were several holiday-themed items that entered the ranking. One notable change was the "PAPER CHAIN KIT 50'S CHRISTMAS" and the "RABBIT NIGHT LIGHT" became the top two items in the fourth quarter. These items experienced a substantial increase in support compared to the previous quarter which indicates a strong demand for festive and seasonal products during the holiday season. The "PAPER CHAIN KIT 50'S CHRISTMAS" made a significant shift in the ranking and jumped to the top spot with a support of 0.1075. This marked a substantial increase in both total sales quantity and total sales value which suggests a high demand for this holiday-themed item during the festive season. Businesses could capitalize on this trend by offering more Christmas-themed items and ensuring sufficient stock levels for the popular items. Conversely, the "JUMBO BAG RETROSPOT" which held the top spot in the previous quarter dropped to the fifth position in the fourth quarter. This decline in ranking could be attributed to the shift in customer preferences towards holiday-themed items during the festive seasons. Moreover, there were several seasonal items that entered the ranking in this quarter. For instance, the items included "JUMBO BAG 50'S CHRISTMAS", "WOODEN STAR CHRISTMAS SCANDINAVIAN", "WOODEN HEART CHRISTMAS SCANDINAVIAN", and "SET OF 20 VINTAGE CHRISTMAS NAPKINS". Businesses can make informed decisions to maximize sales and profitability during key seasonal periods.

Table 4.1.8. Product affinity analysis for 2011 Q4

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0208	0.7150	27.4937	399	2792	8880.52
2	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0208	0.8000	27.4937	399	2792	8880.52
3	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0204	0.6304	24.2428	422	2957	9333.63
4	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0204	0.7838	24.2428	422	2957	9333.63
5	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0226	0.7778	24.0570	445	3107	9900.57
6	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0226	0.7000	24.0570	445	3107	9900.57
7	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0243	0.7936	23.7206	494	2659	3652.55
8	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS	0.0243	0.7269	23.7206	494	2659	3652.55
9	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0212	0.6927	21.5178	482	2726	3781.07
10	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0212	0.6594	21.5178	482	2726	3781.07
12	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0222	0.6900	20.6233	500	2867	3962.12
12	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0222	0.6639	20.6233	500	2867	3962.12
13	CHARLOTTE BAG PINK POLKADOT	RED RETROSPOT CHARLOTTE BAG	0.0204	0.7880	20.5353	469	6742	9502.03
14	RED RETROSPOT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0204	0.5311	20.5353	469	6742	9502.03
15	SET 12 COLOUR PENCILS DOLLY GIRL	SET 12 COLOUR PENCILS SPACEBOY	0.0208	0.6948	19.6935	478	5072	3781.50
16	SET 12 COLOUR PENCILS SPACEBOY	SET 12 COLOUR PENCILS DOLLY GIRL	0.0208	0.5896	19.6935	478	5072	3781.50
17	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOOEDN HEART DECORATIONS	0.0224	0.7294	18.4650	549	3490	4795.83
18	SET OF 3 WOOEDN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0224	0.5758	18.4650	549	3490	4795.83
19	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN HEART DECORATIONS	0.0232	0.6933	17.5515	567	3631	4976.88
20	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0232	0.5872	17.5515	567	3631	4976.88

In this last quarter, there were several new rules appeared in the ranking which mostly involved Christmas-related items such as “SET OF 3 WOODEN TREE DECORATIONS”, “SET OF 3 WOODEN STOCKING DECORATION”, “SET OF 3 WOODEN SLEIGH DECORATIONS”, and “SET OF 3 WOODEN HEART DECORATIONS”. Although there were many new entries in the ranking, however, the top spot of the ranking remained by the rules related to teacup sets such as “GREEN REGENCY TEACUP AND SAUCER” and “PINK REGENCY TEACUP AND SAUCER”. This rule experienced a substantial decrease in support which might be attributed to lowering of interest among customers in this period as well as the total sales value contributed by this rule was relatively lower which only accumulated for 8880.52 compared to the previous quarter. To further boost the sales of the teacup and saucer sets, businesses could cross-promote these sets with other related items such as tea sets, tableware, or kitchen accessories to increase the visibility of the products and attract customers who are interested in complementary items. In terms of new entries, the rule involving “SET OF 3 WOODEN TREE DECORATIONS” and “SET OF 3 WOODEN STOCKING DECORATION” obtained the eighth position in the ranking with a significant lift of 23.7206 which was considered as an outstanding performance for its first try. Similar to the previous quarter, the Charlotte bags remained in their position in the ranking but the consequent changed again. The popularity of the “WOODLAND CHARLOTTE BAG” and “STRAWBERRY CHARLOTTE BAGS” was replaced by the new combination of Charlotte bags which were “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG”. The association between these two items is supported by a high lift of 20.5353 and contributed a total of 9502.03 sales value to the overall revenue. According to this insight, businesses could allocate fewer resources to marketing the Charlotte bags and shift the attention to introducing new designs or styles. It can be observed that the Charlotte bags had consistent popularity across quarters but the popular Charlotte bags were different in design from time to time. Hence, innovation is the key to keeping the product offerings fresh and appealing to customers and businesses could experiment with different colours, patterns, materials, or embellishments to attract new customers and retain existing ones.

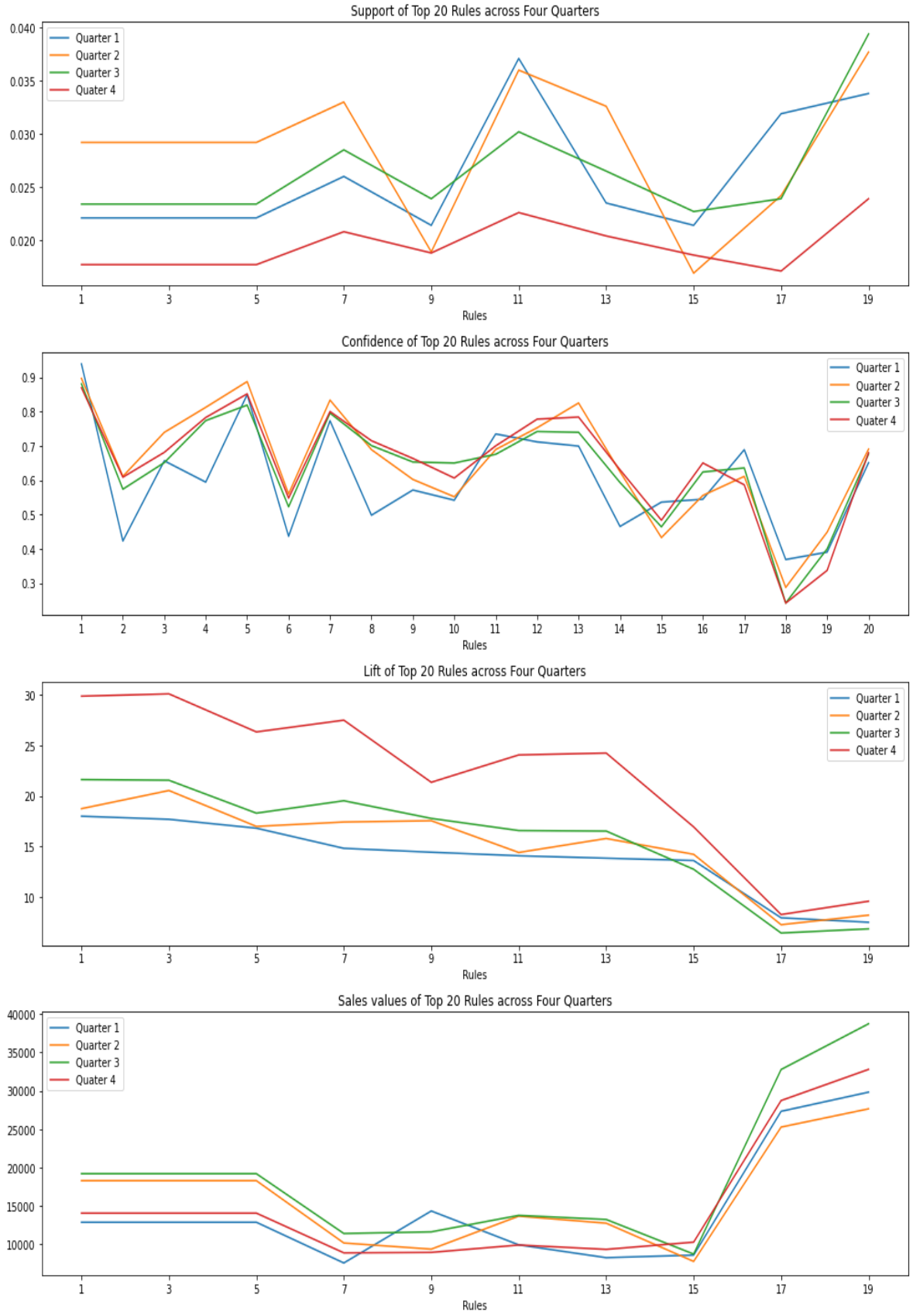


Figure 4.1.5. Metrics fluctuation across quarters

For the quarterly market basket analysis, there are several noteworthy observations to be discussed to discover the hidden trend in the market during the quarterly periods. First of all, Rules 11 and 12 which involved the "ROSES REGENCY TEACUP AND SAUCER" and "GREEN REGENCY TEACUP AND SAUCER" stand out particularly with significant fluctuations in metrics. In the first quarter, this rule exhibited a support of 0.0371, a lift of 14.0899, and a total sales value of 9921.86. However, the rules experienced a notable fluctuation in support and total sales value in the subsequent quarters. The support dropped gradually to 0.0226 while the sales value had its peak value of 13760.93 in the third quarter before experiencing a decline to 9900.57. This insight suggests that there was a reduction in the frequency of transactions involving these items as well as the revenue contributed by these items reduced. The decline in performance could be attributed to several factors such as changing customer preferences in colour variants, the introduction of new competing products, or seasonal variations in demand. Businesses could introduce new variants or designs of these items to generate renewed interest among customers. Similarly, rules 17 and 18 involving "JUMBO BAG BAROQUE BLACK WHITE" and "JUMBO BAG RED RETROSPOT" displayed a decreasing trend in terms of support and lift. The support declined steadily from the first quarter which indicates a potential decrease in the popularity of these items over time. The confidence also showed a downward trend which suggests that the association between these items became weaker over the quarters. This decline could indicate an emergence of alternative product choices that compete with these items. The decrease in the support could be addressed by optimizing the placement of both items in the store and highlighting them in prominent locations to increase visibility and encourage impulse purchases. Moreover, Rules 19 and 20 which involved "JUMBO BAG RED RETROSPOT" and "JUMBO BAG PINK POLKADOT" show a decreasing trend in terms of support and sales value over the quarters. In the fourth quarter, the sales value increased to 32793.21 while conversely the support declined to 0.0239. The total sales value of this set of rules contributed the highest for all four quarters, hence, businesses should properly manage the inventory level for these items to ensure that sufficient stock is maintained to meet customer demand and avoid losing sales opportunities of high-priced items.

4.2 Monthly Market Basket Analysis

On top of quarterly data, the market basket analysis is then conducted on the monthly transaction data of the top three consecutive months with the highest transaction counts in a year. The months involved are October, November, and December. There are several interesting insights and findings into customer behavior and purchasing patterns during the holiday season could be observed. Firstly, a significant increase in transaction count from October to November has been witnessed where the transactions experienced an substantial increase by 31%, from 2637 to 3462, and this showcases the customer activity and willingness to spend surged as the holiday season approaches. This trend could be explained by customers preparing for various festivities such as Thanksgiving, Black Friday, and Christmas which was around the corner. Based on this insight, businesses and retailers could take the opportunity to improve their sales by planning promotions, discounts, and marketing campaigns to attract and retain customers during the peak period. Certain popular products are advised to be stocked up for better storage management and avoid the case of out-of-stock from happening during the peak period. Another valuable insight is that there was a sharp decline in transaction count after November, with transactions decreasing to a mere 1015. This is due to the data provided by the dataset being up until the half of December when the last transaction data was recorded on the 15th of December 2011. However, the decline could also be attributed to other factors such as customer fatigue from excessive spending in the previous two months, completion of gift shopping, and a shift in focus towards other priorities including year-end financial planning. At this moment, businesses should be mindful of this trend and adjust their storage management to mitigate the loss from the decreasing customer activity during this period. In conclusion, based on these insights, businesses could design specific strategies to maximize their profits during the holiday season such as offering targeted promotions and discounts during the period between October and November. The action of price-cutting could attract more price-sensitive customers and enhance customer engagement and loyalty during key shopping events.

Table 4.2.1. Top 20 highest frequency itemsets in the month of October 2011

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PAPER CHAIN KIT 50'S CHRISTMAS	0.0899	237	3189	9419.39
2	JUMBO BAG RED RETROSPOT	0.0770	203	4803	9409.95
3	HOT WATER BOTTLE KEEP CALM	0.0717	189	1315	6066.69
4	REGENCY CAKESTAND 3 TIER	0.0686	181	1027	12660.84
5	WHITE HANGING HEART T-LIGHT HOLDER	0.0656	173	1700	4842.24
6	DOORMAT KEEP CALM AND COME IN	0.0637	168	1330	10111.95
7	SET OF 3 CAKE TINS PANTRY DESIGN	0.0611	161	559	2968.47
8	JUMBO BAG 50'S CHRISTMAS	0.0611	161	2273	4620.59
9	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0607	160	1645	4631.67
10	POPCORN HOLDER	0.0584	154	5865	4778.77
11	ASSORTED COLOUR BIRD ORNAMENT	0.0580	153	2736	4645.92
12	BAKING SET 9 PIECE RETROSPOT	0.0557	147	617	3111.75
13	POSTAGE	0.0550	145	330	7536.78
14	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0546	144	3028	1029.32
15	JUMBO BAG VINTAGE DOILY	0.0546	144	2047	4006.71
16	SET OF 20 VINTAGE CHRISTMAS NAPKINS	0.0538	142	1613	1400.69
17	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0531	140	2354	788.32
18	HAND WARMER OWL DESIGN	0.0527	139	1627	3282.64
19	HOT WATER BOTTLE TEA AND SYMPATHY	0.0523	138	905	5051.51
20	PLAYING CARDS KEEP CALM & CARRY ON	0.0520	137	1401	1727.63

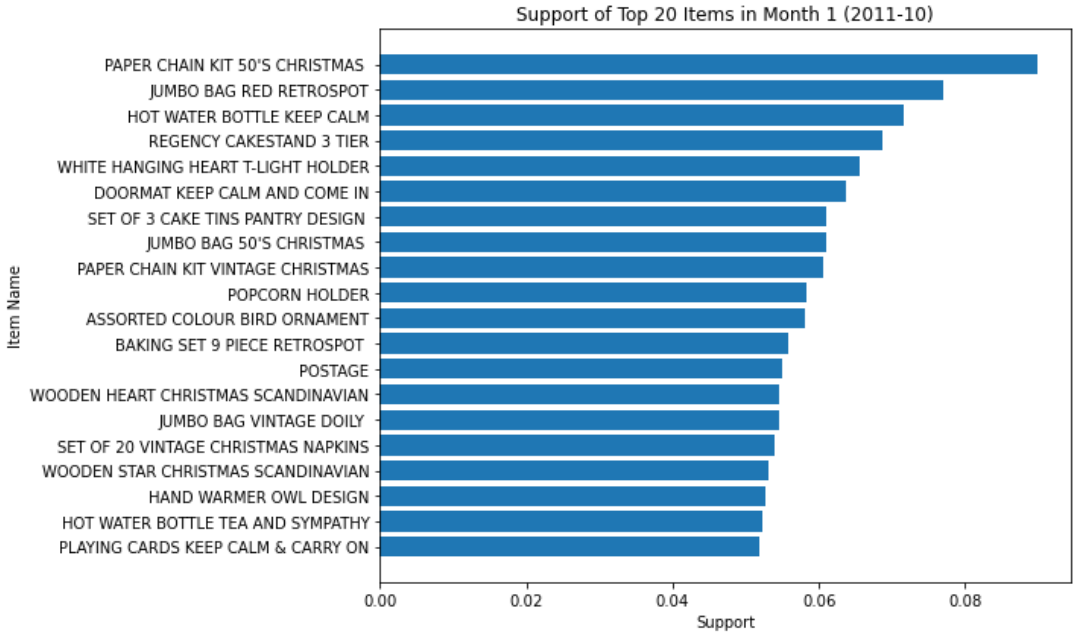


Figure 4.2.1. Visualization of support for top 20 popular items in October 2011

According to the result, "PAPER CHAIN KIT 50'S CHRISTMAS' is recognized as the top-selling item in the month of October with a total of 237 receipts, a total sale quantity of 3189, and a total sales value of 9419.39. The high support value of 0.0899 indicates that it was purchased in nearly 9% of all transactions during the month. Businesses could capitalize on the popularity of this item by promoting it as a must-have and the most trending Christmas decoration with bundle deals to attract more customers. Another notable item is "REGENCY CAKESTAND 3 TIER" which had a relatively high support value of 0.0686 as well. Although the support of this item is not that high compared to the previous Christmas-related decoration, however, it generated the most revenue in the month with a sales value of 12660.84 among the items. Businesses could expand their range of similar products or offer complementary items to enhance the customer's shopping experience. The "DOORMAT KEEP CALM AND COME IN" also stood out with a support value of 0.0637. This item might be viewed as an ordinary household item, but businesses could take inspiration from this and focus on offering unique and functional home decoration items to attract more customers. Similarly, the "HOT WATER BOTTLE KEEP CALM" has a relatively high sales support as well. The businesses could use this information to adjust pricing strategies and gain a chance of a potential increase in profitability.

Table 4.2.2. Product affinity analysis for the month of October 2011

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0201	0.8030	27.8630	147	1261	3732.87
2	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0201	0.6974	27.8630	147	1261	3732.87
3	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0231	0.8026	23.7813	180	1082	1428.29
4	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0231	0.6854	23.7813	180	1082	1428.29
5	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0209	0.8333	23.3777	165	1336	3943.20
6	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0209	0.5851	23.3777	165	1336	3943.20
7	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0239	0.8289	23.2546	176	1359	4101.81
8	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0239	0.6702	23.2546	176	1359	4101.81
9	ROLL WRAP 50'S CHRISTMAS	ROLL WRAP 50'S RED CHRISTMAS	0.0209	0.5500	23.0214	163	1591	1986.96
10	ROLL WRAP 50'S RED CHRISTMAS	ROLL WRAP 50'S CHRISTMAS	0.0209	0.8730	23.0214	163	1591	1986.96
11	SET 12 COLOUR PENCILS DOLLY GIRL	SET 12 COLOUR PENCILS SPACEBOY	0.0235	0.6889	20.1844	185	2261	1534.45
12	SET 12 COLOUR PENCILS SPACEBOY	SET 12 COLOUR PENCILS DOLLY GIRL	0.0235	0.6889	20.1844	185	2261	1534.45
13	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0224	0.6629	19.6418	188	1256	1665.07
14	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0224	0.6629	19.6418	188	1256	1665.07
15	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0220	0.5179	17.9683	211	1386	1806.34
16	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN HEART DECORATIONS	0.0220	0.7632	17.9683	211	1386	1806.34
17	CHARLOTTE BAG PINK POLKADOT	RED RETROSPOT CHARLOTTE BAG	0.0201	0.7794	17.5667	189	2048	2808.93
18	RED RETROSPOT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0201	0.4530	17.5667	189	2048	2808.93
19	WOODEN HEART CHRISTMAS SCANDINAVIAN, WOODEN TREE CHRISTMAS SCANDINAVIAN	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0288	0.9048	17.2888	396	7089	2409.11
20	WOODEN STAR CHRISTMAS SCANDINAVIAN	WOODEN HEART CHRISTMAS SCANDINAVIAN, WOODEN TREE CHRISTMAS SCANDINAVIAN	0.0288	0.5507	17.2888	396	7089	2409.11

The product affinity analysis for the first month which is October 2011 provides several significant insights in terms of the association rule's metrics such as support, lift, and confidence. First of all, the rule involved "PINK REGENCY TEACUP AND SAUCER" and "GREEN REGENCY TEACUP AND SAUCER" had a support of 0.201 which implies that around 20% of the monthly transactions include both items. The good performance of this rule is also attributed to the high confidence of 0.8030 and a lift value of 27.863. As suggestions, businesses could leverage this association by introducing promotions that bundle these items together or offering discounts for purchasing both. Additionally, these items could be strategically placed near each other in the store to encourage cross-selling and improve customer experience by targeting customers interested in regency-style tea sets who already decided to purchase them together. There were also other regency-style tea sets that appeared in the ranking but in different colours such as "ROSES REGENCY TEACUP AND SACUER" which had an association with tea sets in pink and green colour. Businesses should be more concerned about colour assortments of their products as they can effectively influence buying decisions and customer behavior. In addition, there was a strong association between "SET OF 3 WOODEN TREE DECORATIONS" and "SET OF 3 WOODEN STOCKING DECORATION" which is verified by the high lift of 23.7813. Businesses should take note of this association as it reveals the trend of Christmas shopping is approaching. By featuring them in themed holiday displays and emphasizing the complementary nature of these items, businesses could boost the average transaction value and drive additional sales. Similarly, there were other Christmas-related items having a strong association among them. For instance, "WOODEN TREE CHRISTMAS SCANDINAVIAN", "WOODEN HEART CHRISTMAS SCANDINAVIAN", and "WOODEN STAR CHRISTMAS SCANDINAVIAN" showed up in the ranking of top association rules in October with a very strong confidence of 0.9048 and a support of 0.0288. Businesses could leverage this association between these items by creating holiday-themed gift sets that include all three items and featuring them prominently in holiday marketing campaigns. They should also capitalize on this opportunity to promote other similar Christmas-related items to appeal to customers' desire for cohesive and festive decorations and eventually drive higher sales.

Table 4.2.3. Top 20 highest frequency itemsets in the month of November 2011

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0.1511	523	14883	34422.09
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.1176	407	7868	28883.04
3	HOT WATER BOTTLE KEEP CALM	0.0887	307	2285	11500.87
4	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0823	285	3786	12831.54
5	JUMBO BAG 50'S CHRISTMAS	0.0797	276	3971	10913.10
6	WHITE HANGING HEART T-LIGHT HOLDER	0.0789	273	4938	14345.22
7	POPCORN HOLDER	0.0774	268	12159	13968.74
8	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0739	256	5364	1787.21
9	JUMBO BAG RED RETROSPOT	0.0731	253	5807	11458.21
10	JUMBO BAG PAISLEY PARK	0.0719	249	3072	7830.34
11	GARDENERS KNEELING PAD KEEP CALM	0.0690	239	2763	5863.75
12	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0685	237	5003	1658.29
13	HAND WARMER OWL DESIGN	0.0679	235	2530	5756.59
14	LUNCH BAG PAISLEY PARK	0.0676	234	2325	3758.73
15	JUMBO BAG VINTAGE DOILY	0.0673	233	3602	8996.75
16	BAKING SET 9 PIECE RETROSPOT	0.0653	226	1139	5651.21
17	REGENCY CAKESTAND 3 TIER	0.0635	220	968	12798.48
18	CHOCOLATE HOT WATER BOTTLE	0.0633	219	1507	7676.50
19	HOT WATER BOTTLE TEA AND SYMPATHY	0.0621	215	1746	10012.64
20	ASSORTED COLOUR BIRD ORNAMENT	0.0612	212	5254	8440.94

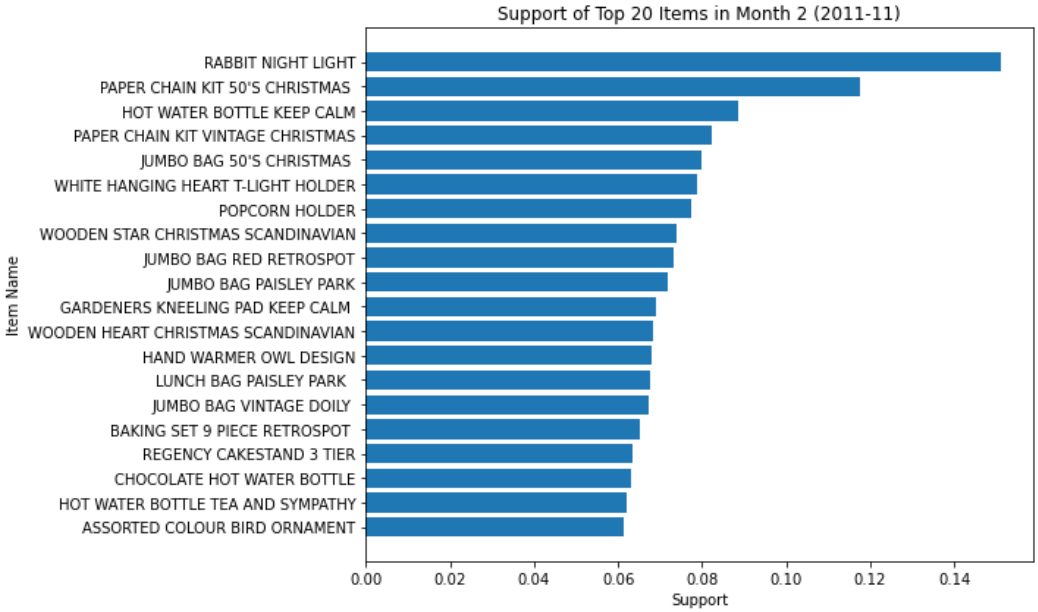


Figure 4.2.2. Visualization of support for top 20 popular items in November 2011

In comparing the most 20 popular items between October and November, certain significant fluctuations in terms of support, total sales quantity, and total sales value could be observed. For instance, the "RABBIT NIGHT LIGHT" obtained its entry to the ranking with an extremely high total sales value of 34422.09 and a support value of 0.1511. This significant surge in demand suggests businesses take note of such a trend and consider stocking up on such popular items to meet the increased demand. Another notable change is the fluctuation in support value for certain items. The "PAPER CHAIN KIT 50'S CHRISTMAS" experienced a surge in support from 0.0899 in October to 0.1176 in November which is around a 30% increase and this shift in support indicates a change in the association of this item with other products in the market basket. Businesses should closely monitor support values for their products to identify any emerging trends and adjust their marketing strategies accordingly. Furthermore, there are other new entries and exits of items in the ranking aside from the "RABBIT NIGHT LIGHT". For instance, "GARDENERS KNEELING PAD KEEP CALM" entered the top 20 list in November which reflects a sudden increase in demand or popularity for this item. Conversely, "SET OF 3 CAKE TINS PANTRY DESIGN" exited the list which suggests a decline in demand for this particular product. Businesses should analyze these changes to understand emerging trends and adjust their storage management for related products to minimize loss.

Table 4.2.4. Product affinity analysis for the month of November 2011

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0202	0.8046	28.1365	188	961	3270.71
2	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0202	0.7071	28.1365	188	961	3270.71
3	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN HEART DECORATIONS	0.0217	0.6410	26.4194	412	2398	3360.38
4	SET OF 3 WOODEN HEART DECORATION, SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0217	0.8929	26.4194	412	2398	3360.38
5	SET OF 3 WOODEN HEART DECORATION, SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0217	0.9036	26.0693	412	2398	3360.38
6	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN HEART DECORATION, SET OF 3 WOODEN TREE DECORATIONS	0.0217	0.6250	26.0693	412	2398	3360.38
7	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0202	0.7071	23.9988	203	1118	3741.30
8	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0202	0.6863	23.9988	203	1118	3741.30
9	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0260	0.7692	22.1923	256	1285	1797.55
10	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0260	0.7500	22.1923	256	1285	1797.55
11	CHARLOTTE BAG PINK POLKADOT	RED RETROSPOT CHARLOTTE BAG	0.0217	0.7732	21.4144	229	3581	4738.54
12	RED RETROSPOT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0217	0.6000	21.4144	229	3581	4738.54
13	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATION	0.0240	0.7217	21.3561	247	1276	1803.24
14	SET OF 3 WOODEN TREE DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0240	0.7094	21.3561	247	1276	1803.24
15	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0234	0.7043	20.3204	255	1275	1806.87
16	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0234	0.6750	20.3204	255	1275	1806.87
17	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	0.0217	0.5245	20.1748	412	2398	3360.38
18	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN HEART DECORATIONS	0.0217	0.8333	20.1748	412	2398	3360.38
19	SET OF 12 COLOUR PENCILS DOLLY GIRL	SET OF 12 COLOUR PENCILS SPACEBOY	0.0208	0.6857	18.4027	242	2369	1926.75
20	SET OF 12 COLOUR PENCILS SPACEBOY	SET OF 12 COLOUR PENCILS DOLLY GIRL	0.0208	0.5581	18.4027	242	2369	1926.75

By comparing the product affinity analysis of October and November, it can be observed that the rule involved “PINK REGENCY TEACUP AND SAUCER” and “GREEN REGENCY TEACUP AND SAUCER” had a stable support level between the two months, however, there was a slight decrease in total sales value from 3732.87 in the previous month to 3270.71 which is a roughly 12.34% of decrease. To address this, businesses could consider reevaluating their pricing strategy, offering bundle deals or promotions, or enhancing their marketing efforts to maintain or increase sales. Furthermore, the support of the rule involving “SET OF 3 WOODEN TREE DECORATIONS” and “SET OF 3 WOODEN STOCKING DECORATION” remained relatively stable while the sales value drastically increased from 1428.29 in October to 1797.55 which reflects the association between these items might be getting weaker in this month. To maintain the popularity and sales of the items, businesses should allocate more resources to promote these items or offer special promotions to encourage bundled purchases. Another notable observation is that the rule including “SET 12 COLOUR PENCILS DOLLY GIRL” and “SET 12 COLOUR PENCILS SPASCEBOY” dropped drastically in terms of ranking from the eleventh position to nineteenth which nearly dropped out of the ranking. This decline in ranking was caused by the high competitiveness of other products in the ranking while the rules related to Christmas-related items were getting more popular and having stronger association in this month. Moreover, there was a noteworthy rule that involved “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG” which ranked eleventh in November and had the highest sales value of 4738.52 among the rules in the ranking. The approximately 70% increase in the sales value made it outperform other rules which had higher lift and support. By noticing the performance of these two items was getting better, businesses could continue to market these bags as a complementary pair and appeal to customers looking for matching or coordinated accessories. Aside from the fluctuation of rule’s metrics, there were certain rules exited the ranking this month. For instance, the rule with the items “ROLL WRAP 50’S RED CHRISTMAS” and “ROLL WRAP 50’S CHRISTMAS” which had its position in the top 10 rules ranking in the previous month dropped out of the ranking in this month.

Table 4.2.5. Top 20 highest frequency itemsets in the month of December 2011

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0.1369	139	4298	9618.01
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.1192	121	2089	6870.71
3	HOT WATER BOTTLE KEEP CALM	0.0916	93	680	3974.68
4	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0887	90	862	2763.56
5	HAND WARMER OWL DESIGN	0.0808	82	812	1756.11
6	JUMBO BAG 50'S CHRISTMAS	0.0739	75	1075	2224.15
7	BICYCLE PUNCTURE REPAIR KIT	0.0729	74	1240	2595.79
8	GARDENERS KNEELING PAD KEEP CALM	0.0719	73	699	1649.27
9	POPCORN HOLDER	0.0709	72	6198	5409.86
10	CHOCOLATE HOT WATER BOTTLE	0.0680	69	575	3470.24
11	REGENCY CAKESTAND 3 TIER	0.0650	66	441	5902.92
12	HAND WARMER RED LOVE HEART	0.0631	64	474	1011.64
13	HOT WATER BOTTLE TEA AND SYMPATHY	0.0631	64	608	3734.80
14	WHITE HANGING HEART T-LIGHT HOLDER	0.0621	63	777	2204.07
15	SCOTTIE DOG HOT WATER BOTTLE	0.0611	62	403	2120.11
16	ALARM CLOCK BAKELIKE RED	0.0601	61	1060	3814.40
17	HAND WARMER BIRD DESIGN	0.0591	60	743	1537.48
18	SET OF 20 VINTAGE CHRISTMAS NAPKINS	0.0581	59	705	632.01
19	GARDENERS KNEELING PAD CUP OF TEA	0.0581	59	441	818.57
20	JUMBO BAG RED RETROSPOT	0.0571	58	1090	2195.70

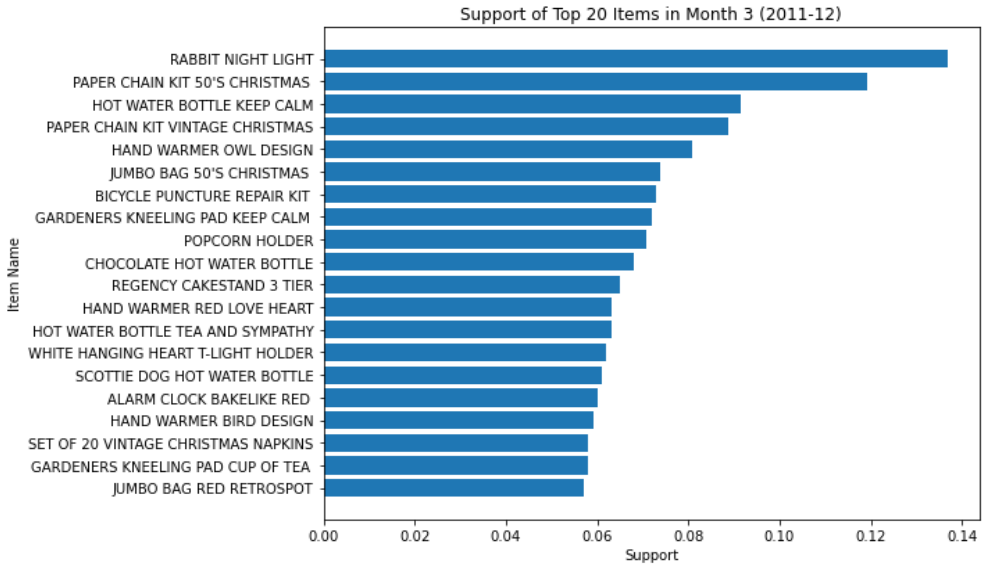


Figure 4.2.3. Visualization of support for top 20 popular items in December 2011

According to the most 20 popular items of December, valuable insights into evolving customer preferences and purchasing behaviors during the holiday season could be observed. One significant change observed is the fluctuations in sales quantity and sales value for certain items. For example, the “RABBIT NIGHT LIGHT” maintained its top position as the top-selling item in December, however, there was a decrease in total sales quantity and total sales value by approximately 71.13% and 72.05% respectively compared to the previous month. Businesses should precisely adjust their inventory levels to accommodate these fluctuations in demand as the “RABBIT NIGHT LIGHT” experienced an increase in demand in the previous month but its demand declined in subsequent months. Similarly, the “PAPER CHAIN KIT 50’S CHRISTMAS” experienced a decrease in total sales quantity and total sales value by approximately 73.52% and 76.18% respectively, in December compared to November. Conversely, “HAND WARMER OWL DESIGN” saw an increase in both total sales quantity and total sales value by doubling up. This might be attributed to the weather getting colder as the winter was approaching. Moreover, there were new entries and exits in the list as well. “BICYCLE PUNTURE REPAIR KIT” entered the list and “JUMBO BAG PAISLEY PARK” exited the top 20 list. The overall sales value of the top items in the ranking is relatively low compared to the previous two months which suggests businesses have to adjust their product offerings and marketing strategies accordingly.

Table 4.2.6. Product affinity analysis for the month of December 2011

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0217	0.8800	30.8000	58	292	426.71
2	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0217	0.7586	30.8000	58	292	426.71
3	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0227	0.8846	28.0589	98	884	2921.72
4	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0227	0.7188	28.0589	98	884	2921.72
5	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0227	0.9200	27.4647	98	884	2921.72
6	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0227	0.6765	27.4617	98	884	2921.72
7	CHARLIE + LOLA PINK HOT WATER BOTTLE	CHARLIE + LOLA RED HOT WATER BOTTLE	0.0227	0.8519	26.2009	61	470	653.14
8	CHARLIE + LOLA RED HOT WATER BOTTLE	CHARLIE+LOLA PINK HOT WATER BOTTLE	0.0227	0.6970	26.2009	61	470	653.14
9	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0276	0.8750	26.1213	66	630	2057.46
10	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0276	0.8235	26.1213	66	630	2057.46
11	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0227	0.8214	26.0547	98	884	2921.72
12	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, GREEN REGENCY TEACUP AND SAUCER	0.0227	0.7188	26.0547	98	884	2921.72
13	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0246	0.7813	24.7803	64	570	1876.94
14	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0246	0.7813	24.7803	64	570	1876.94
15	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0256	0.8125	24.2555	66	568	1909.04
16	ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0256	0.7647	24.2555	66	568	1909.04
17	CERAMIC CHERRY CAKE MONEY BANK	CERAMIC HEART FAIRY CAKE MONEY BANK	0.0207	0.7000	22.9194	63	886	469.92
18	CERAMIC HEART FAIRY CAKE MONEY BANK	CERAMIC CHERRY CAKE MONEY BANK	0.0207	0.6774	22.9194	63	886	469.92
19	DOLLY GIRL LUNCH BOX	SPACEBOY LUNCH BOX	0.0207	0.7778	22.5556	67	616	1184.33
20	SPACEBOY LUNCH BOX	DOLLY GIRL LUNCH BOX	0.0207	0.6000	22.5556	67	616	1184.33

In the comparison of the product affinity analysis for November and December, there were several notable observations in terms of metrics fluctuation and shifts in ranking. First of all, the top spot of the ranking which was previously dominated by the teacup sets, “PINK REGENCY TEACUP AND SAUCER” and “GREEN REGENCY TEACUP AND SAUCER” was replaced by the rule including the items “SET OF 3 WOODEN TREE DECORATIONS” and “SET OF 3 WOODEN STOCKING DECORATION” with a significantly high support value of 0.0217. Although this rule shifted in ranking, however, it had a decline in terms of the metrics of support and sales value. The support for this rule significantly decreased from 0.026 to 0.0217 which is a decrease of nearly 16.5%. Businesses could capitalize on this trend of offering wooden decorations items as part of a larger holiday decoration collection which will allow customers to create cohesive and festive displays in their homes. The items could also be bundled together at a discounted price or offering them as part of a “buy one, get one half off” promotion to encourage cross-selling. For the regency teacup sets, the rule involving “PINK REGENCY TEACUP AND SAUCER” and “GREEN REGENCY TEACUP AND SAUCER” dropped from the top spot for the ranking to fourteenth in December despite the support for this rule increased by 21.7% to 0.0246. However, the association rules for the pink, green, and roses teacup sets remained their position in the ranking which indicates there were still strong associations between these products. Businesses might consider continuing to promote these sets as ideal gifts for tea enthusiasts or collectors to maintain or increase their sales for the upcoming period. In terms of new entries of the ranking, there are two sets of rules entered the ranking this month which are “CERAMIC CHERRY CAKE MONEY BANK”, “CERAMIC HEART FAIRY CAKE MONEY BANK”, “DOLLY GIRL LUNCH BOX”, and “SPACEBOY LUNCH BOY”. Despite these sets of rules just entering the ranking, this trend might indicate a strong appeal to parents who started to look for fun and practical lunchtime solutions for their children. Hence, businesses could consider creating more themed lunch sets that include these boxes along with matching accessories such as water bottles and snack containers. They could also be prepared to launch the “back-to-school” promotion to attract families preparing for the new school year.

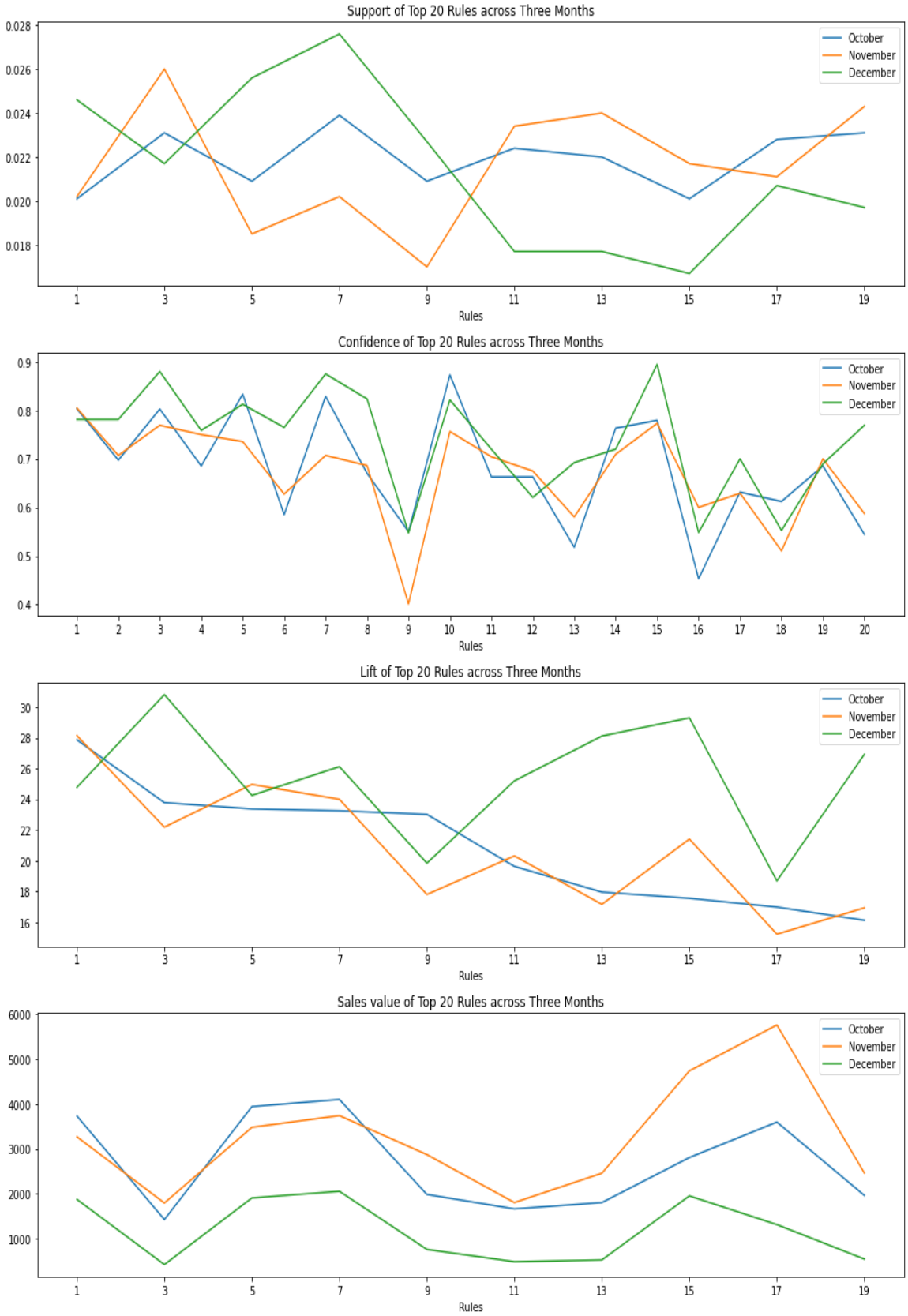


Figure 4.2.4. Metrics fluctuation across three months

By analyzing the fluctuation of metrics for the top 20 rules across three three-month periods, a clearer understanding of the trends could be obtained. Significant variations in support, lift, confidence, and sales value metrics indicate the degree of change in the performance of each rule over the specified period. First of all, Rule 1 and 2 which involved the "PINK REGENCY TEACUP AND SAUCER" and "GREEN REGENCY TEACUP AND SAUCER" exhibits a consistent increase in support over the three months which indicates a growing association between these items. There was an average increase of approximately 0.0015 in support from October to December while the sales value constantly decreased slightly from October to December. However, there is a more substantial decline in sales value with a decrease of approximately 1865.93 over the same period. The decrease in sales value despite stable support and lift metrics suggests a potential issue with the pricing or promotion strategy. It may be beneficial to review pricing competitiveness and consider targeted promotions to stimulate sales. Furthermore, Rules 3 and 4 which involved the itemsets "SET OF 3 WOODEN TREE DECORATIONS" and "SET OF 3 WOODEN STOCKING DECORATION" stand out with significant fluctuations in the metrics. In October, the rules had a support of 0.0231 but it dropped to 0.0217 before increasing to 0.0260 in December. This could be due to seasonal factors such as changes in decoration trends during the holiday season or promotional activities that influenced customer purchasing behavior. To address this, businesses could conduct customer surveys or analyze sales data from previous years could provide insights into seasonal trends or changes in customer preferences. Furthermore, Rules 15 and 16 involving "CHARLOTTE BAG PINK POLKADOT" and "RED RETROSPOT CHARLOTTE BAG" experienced a significant increase in sales value which increased from 2809.93 in October to 4738.54 in November but eventually dropped to 1954.56 in December. This indicates the association between the items was fluctuating significantly from time to time. Similarly, the support for the rules was relatively low during the months which eventually contributed to a lower total sales value. Businesses should create promotional bundles or discounts for customers purchasing both items to encourage them to buy both items together to increase the sales quantity and sales value for this rule.

4.3 Biweekly Market Basket Analysis

The market basket analysis is eventually conducted for the biweekly period and the biweekly transaction data is obtained from the top three consecutive months with the highest number of transactions which are October, November, and December. The first biweek which started from the end of September to the beginning of October witnessed 661 transactions. The number of transactions in this period is considered relatively low due to customers adjusting their shopping habits to accommodate seasonal changes and upcoming holidays. Businesses could make use of this opportunity to introduce new seasonal products and promotions to capture early shoppers and get prepared for the upcoming holiday season. The second and third biweek which started from 8 October to 21 October and 22 October to 4 November respectively underwent a significant increase in transaction count. The number of transactions of these two biweeks came to a thousand or more, which might be attributed to the seasonal festive, Halloween. In fact, it is a popular shopping event for costumes, decorations, and party supplies. Businesses could pay more attention to the themed products in order to gain the maximum profits from Halloween enthusiasts and ensure there are sufficient products for the increased demand. Subsequently, the fourth biweek marks a continued growth of customer activity as it accumulated up to 1674 transactions, which is the highest number of transactions among the six biweeks. This period represents a precious opportunity for businesses to maximize sales and revenue due to the period coincides with major shopping events like Black Friday and Cyber Monday. For the last two second biweeks, the number of transactions declined from 1674 to 1611 and eventually ended up with 704 transactions in the last biweek. Businesses could use this time to evaluate their performance during the holiday season, analyze customer trends, and prepare for the season-end clearance sales to clear out the excess inventory. Post-holiday sales and promotions could be introduced to attract bargain-seeking shoppers for businesses to maximize their revenue during the season.

Table 4.3.1. Top 20 highest frequency itemsets in Biweek 1 (24 September 2011 – 7 October 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	JUMBO BAG RED RETROSPOT	0.0832	55	2489	4609.33
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.0802	53	931	2588.93
3	DOORMAT KEEP CALM AND COME IN	0.0772	51	695	5046.24
4	PLAYING CARDS KEEP CALM & CARRY ON	0.0711	47	445	558.67
5	LUNCH BAG SPACEBOY DESIGN	0.0666	44	288	521.53
6	HOT WATER BOTTLE KEEP CALM	0.0651	43	564	2509.64
7	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0651	43	519	1403.77
8	LUNCH BAG VINTAGE DOILY	0.0635	42	322	536.46
9	REGENCY CAKESTAND 3 TIER	0.0635	42	393	4715.58
10	JUMBO BAG VINTAGE DOILY	0.0635	42	752	1423.26
11	ASSORTED COLOUR BIRD ORNAMENT	0.0620	41	588	993.72
12	JUMBO BAG 50'S CHRISTMAS	0.0620	41	466	954.63
13	SPOTTY BUNTING	0.0605	40	178	969.58
14	PARTY BUNTING	0.0605	40	152	1020.88
15	LARGE WHITE HEART OF WICKER	0.0590	39	292	718.98
16	SET OF 3 CAKE TINS PANTRY DESIGN	0.0590	39	95	483.35
17	POPCORN HOLDER	0.0590	39	741	660.78
18	JUMBO BAG VINTAGE CHRISTMAS	0.0560	37	232	490.76
19	LOVE HOT WATER BOTTLE	0.0560	37	138	797.10
20	WHITE HANGING HEART T-LIGHT HOLDER	0.0560	37	159	486.09

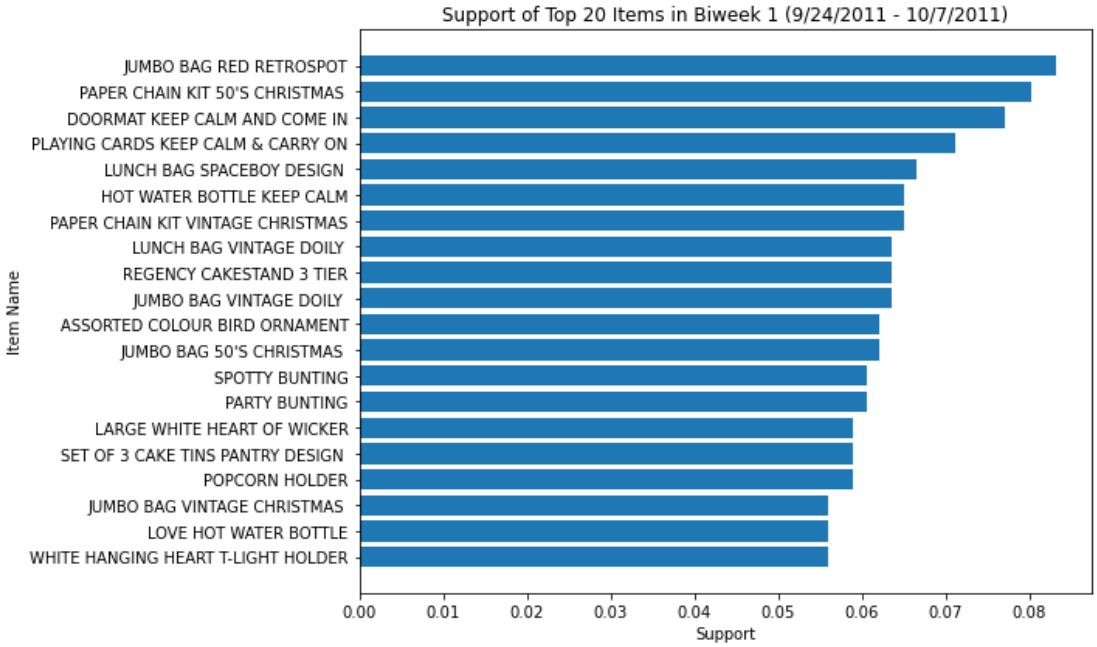


Figure 4.3.1. Visualization of support for top 20 popular items in Biweek 1

The top 20 items in the first biweek reveal several interesting insights for businesses and one notable observation is the presence of seasonal and themed items during the period between the end of September and the beginning of October. Items such as “PAPER CHAIN KIT 50’S CHRISTMAS” and “PAPER CHAIN KIT VINTAGE CHRISTMAS” ranked second and seventh respectively as these items likely experienced increased demand due to the upcoming Christmas season. This insight suggests businesses anticipate and prepare for seasonal trends by stocking up on holiday-themed products. Another interesting observation is the performance of “DOORMAT KEEP CALM AND COME IN” which ranked third with a support value of 0.0772. The high support indicates a strong affinity among customers and businesses could consider offering more products with similar designs to attract more customers. On the other hand, certain items like “LUNCH BAG SPACEBOY DESIGN” and “LUNCH BAG VINTAGE DOILY” ranked lower in terms of support, however, they had relatively high total sales volume compared to other items in the ranking. For instance, the “LUNCH BAG SPACEBOY DESIGN” had a support value of 0.0666 but a total sales value of 521.53 which indicates this item may not be popular among customers overall, but they still attract customers willing to spend more. Businesses could target niche markets or offer premium versions of these products to cater to customers who willing to pay a premium.

Table 4.3.2. Product affinity analysis for Biweek 1 (24 September 2011 – 7 October 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	GLASS APOTHECARY BOTTLE TONIC	GLASS APOTHECARY BOTTLE PERFUME	0.0212	1.0000	34.7895	37	124	484.82
2	FELTCRAFT PRINCESS CHARLOTTE DOLL, FELTCRAFT PRINCESS LOLA DOLL	FELTCRAFT PRINCESS OLIVIA DOLL	0.0212	1.0000	34.7895	58	361	1305.62
3	GLASS APOTHECARY BOTTLE PERFUME	GLASS APOTHECARY BOTTLE TONIC	0.0212	0.7368	34.7895	37	124	484.82
4	FELTCRAFT PRINCESS OLIVIA DOLL	FELTCRAFT PRINCESS CHARLOTTE DOLL, FELTCRAFT PRINCESS LOLA DOLL	0.0212	0.7368	34.7895	58	361	1305.62
5	FELTCRAFT PRINCESS CHARLOTTE DOLL, FELTCRAFT PRINCESS OLIVIA DOLL	FELTCRAFT PRINCESS LOLA DOLL	0.0212	0.8750	32.1319	58	361	1305.62
6	FELTCRAFT PRINCESS LOLA DOLL	FELTCRAFT PRINCESS CHARLOTTE DOLL, FELTCRAFT PRINCESS OLIVIA DOLL	0.0212	0.7778	32.1319	58	361	1305.62
7	FELTCRAFT PRINCESS OLIVIA DOLL	FELTCRAFT PRINCESS LOLA DOLL	0.0242	0.8421	30.9240	37	226	816.85
8	FELTCRAFT PRINCESS LOLA DOLL	FELTCRAFT PRINCESS OLIVIA DOLL	0.0242	0.8889	30.9240	37	226	816.85
9	FELTCRAFT PRINCESS LOLA DOLL, FELTCRAFT PRINCESS OLIVIA DOLL	FELTCRAFT PRINCESS CHARLOTTE DOLL	0.0212	0.8750	27.5417	58	361	1305.62
10	FELTCRAFT PRINCESS CHARLOTTE DOLL	FELTCRAFT PRINCESS LOLA DOLL, FELTCRAFT PRINCESS OLIVIA DOLL	0.0212	0.6667	27.5417	58	361	1305.62
11	FELTCRAFT PRINCESS OLIVIA DOLL	FELTCRAFT PRINCESS CHARLOTTE DOLL	0.0242	0.8421	26.5063	40	253	914.19
12	FELTCRAFT PRINCESS CHARLOTTE DOLL	FELTCRAFT PRINCESS OLIVIA DOLL	0.0242	0.7619	26.5063	40	253	914.19
13	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0257	0.7391	25.7140	43	680	1850.52
14	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0257	0.8947	25.7140	43	680	1850.52
15	FELTCRAFT PRINCESS LOLA DOLL	FELTCRAFT PRINCESS CHARLOTTE DOLL	0.0212	0.7778	24.4815	39	243	880.40
16	FELTCRAFT PRINCESS CHARLOTTE DOLL	FELTCRAFT PRINCESS LOLA DOLL	0.0212	0.6667	24.4815	39	243	880.40
17	WOODEN TREE CHRISTMAS SCANDINAVIAN	WOOEDEN STAR CHRISTMAS SCANDINAVIAN	0.0242	0.8000	22.9913	43	787	267.65
18	WOODEN STAR CHRISTMAS SCANDINAVIAN	WOODEN TREE CHRISTMAS SCANDINAVIAN	0.0242	0.6957	22.9913	43	787	267.65
19	RIBBON REEL MAKING SNOWMEN	RIBBON REEL SNOWY VILLAGE	0.0212	0.7000	21.0318	45	366	615.42
20	RIBBON REEL SNOWY VILLAGE	RIBBON REEL MAKING SNOWMEN	0.0212	0.6364	21.0318	45	366	615.42

Based on the product affinity analysis for the first biweek, several significant association rules stood out and one notable association was between the “GLASS APOTHECARY BOTTLE TONIC” and “GLASS APOTHECARY BOTTLE PERFUME” with a support of 0.0212. This rule had a confidence value of 1.000 which suggests that whenever customers bought the “GLASS APOTHECARY BOTTLE TONIC”, they always bought the “GLASS APOTHECARY BOTTLE PERFUME” as well. Additionally, this rule dominated the top spot of the ranking with a lift value of 34.7895 which indicates that this association is significantly higher than would be expected by chance and making it a strong rule. Businesses could capitalize on the perfect confidence of this rule to bundle these two items together at a slightly discounted price or offer a free item such as a small perfume sample when purchasing both items together to encourage more sales. Another noteworthy association is among the “FELTCRAFT PRINCESS LOLA DOLL”, “FELTCRAFT PRINCESS CHARLOTTE DOLL”, and “FELTCRAFT PRINCESS OLIVLA DOLL”. These dolls were frequently purchased together and it was verified by the high support value of 0.0212 and a perfect confidence of 1.0000 as well. Businesses could consider marketing them as a set or a collection with special packaging or a discount when purchased together. This strategy could possibly increase the perceived value of the dolls as a set and encourage customers to buy all three to boost the total sales value Aside from the rule’s metrics, the rule involving “ROSES REGENCY TEACUP AND SAUCER” and “GREEN REGENCY TEACUP AND SAUCER” had the highest total sales value despite its support and lift were not that high as other rules in the ranking. Since these teacup and saucer sets are frequently purchased together, businesses could consider creating a themed display featuring both sets to enhance the visual appeal of the products and encourage customers to purchase both sets together. Moreover, there was another noteworthy rule which involved “WOODEN TREE CHRISTMAS SCANDINAVIAN” and “WOODEN STAR CHRISTMAS SCANDINAVIAN” with a high confidence and lift. The show up of this rule might indicate the beginning of Christmas for the upcoming period and suggest businesses start stocking up on similar items to maximize revenue. A holiday-themed promotion would be suggested to feature both items and bundle them together at a discounted price or offer a free gift-wrapping service for purchases of both items.

Table 4.3.3. Top 15 highest frequency itemsets in Biweek 2 (8 October 2011 – 21 October 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PAPER CHAIN KIT 50'S CHRISTMAS	0. 0941	106	1508	4399.32
2	JUMBO BAG RED RETROSPOT	0.0825	93	1703	3409.39
3	REGENCY CAKESTAND 3 TIER	0.0816	92	454	5615.73
4	HOT WATER BOTTLE KEEP CALM	0.0799	90	443	2112.29
5	SET OF 3 CAKE TINS PANTRY DESIGN	0.0665	75	337	1739.67
6	JUMBO BAG 50'S CHRISTMAS	0.0657	74	844	1848.37
7	DOORMAT KEEP CALM AND COME IN	0.0630	71	450	3450.38
8	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0630	71	1240	412.68
9	WHITE HANGING HEART T-LIGHT HOLDER	0.0630	71	879	2483.65
10	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0621	70	1516	493.76
11	POPCORN HOLDER	0.0612	69	2159	1830.80
12	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0603	68	754	2092.90
13	JUMBO BAG VINTAGE DOILY	0.0603	68	669	1345.82
14	BAKING SET 9 PIECE RETROSPOT	0.0594	67	241	1241.11
15	RETROSPOT TEA SET CERAMIC 11 PC	0..0586	66	344	1763.36
16	SET OF 6 RIBBONS VINTAGE CHRISTMAS	0.0577	65	441	1320.89
17	SET OF 3 REGENCY CAKE TINS	0.0577	65	448	2201.68
18	POSTAGE	0.0568	64	156	3338.32
19	SET OF 20 VINTAGE CHRISTMAS NAPKINS	0.0568	64	804	682.62
20	LUNCH BAG RED RETROSPOT	0.0550	62	685	1356.53

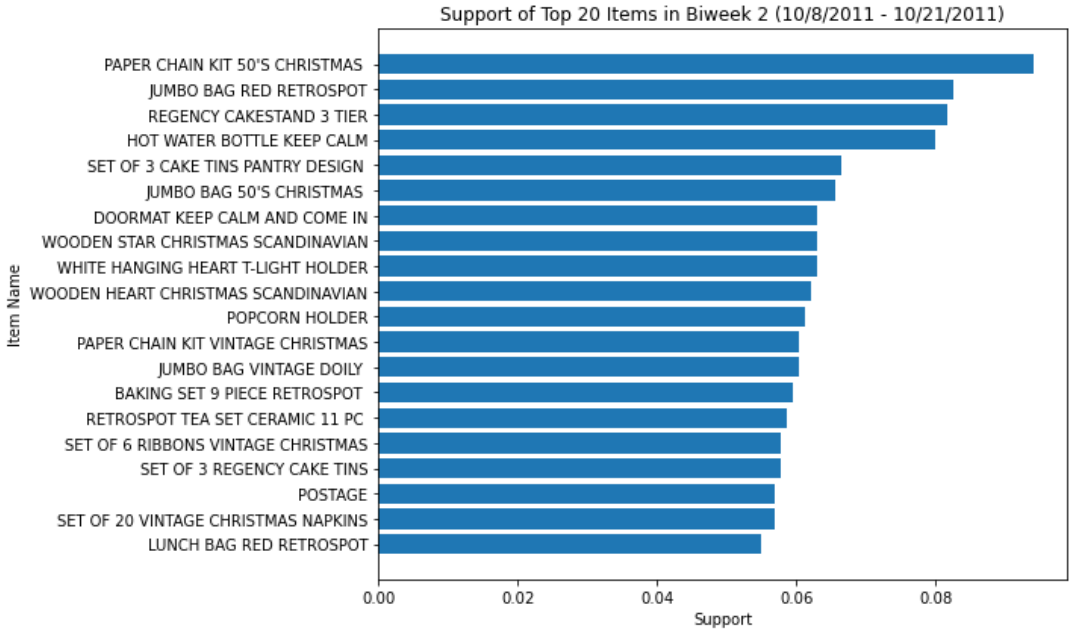


Figure 4.3.2. Visualization of support for top 20 popular items in Biweek 2

In the second biweek, the "PAPER CHAIN KIT 50'S CHRISTMAS" item witnessed a substantial increase in support from 0.0802 in the first biweek to 0.0941 in the second biweek. The approximately 17.4% of increase suggests a higher proportion of transactions involving this item in the second biweek. The continued growth in popularity for this item indicates a potential surge in demand for Christmas-themed products for the following period. Bundling this item with other Christmas-themed products or offering discounts for bulk purchases could further drive sales. Conversely, the "JUMBO RED RETROSPOT" which ranked first in the previous biweek experienced a decrease in total sales quantity from 2489 to 1703 units. Despite there was a nearly 31.5% decrease in total sales quantity, the support for this item remained relatively stable which reflects a consistent level of customer interest during this period. Businesses should reassess their inventory levels and adjust orders accordingly to avoid overstocking for this item. They could also consider offering promotions or incentives to stimulate demand for these items. Additionally, the "REGENCY CAKESTAND 3 TIER" experienced a slight rise in total sales quantity from 393 in the first biweek to 454 in the second biweek. On top of that, the total sales value for this item increased by 19.1% to 5615.73 as well. Businesses could leverage this insight by promoting complementary items or upselling higher-priced products to increase their average transaction value.

Table 4.3.4. Product affinity analysis for Biweek 2 (8 October 2011 – 21 October 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	REGENCY TEA PLATE GREEN	REGENCY TEA PLATE ROSES	0.0213	0.8889	31.3056	59	686	1156.06
2	REGENCY TEA PLATE ROSES	REGENCY TEA PLATE GREEN	0.0213	0.7500	31.3056	59	686	1156.06
3	GREEN REGENCY TEACUP AND SAUCER, ROSES REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0213	0.8276	28.2633	122	780	2414.72
4	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, ROSES REGENCY TEACUP AND SAUCER	0.0213	0.7273	28.2633	122	780	2414.72
5	PICK REGENCY TEACUP AND SAUCER, ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0213	0.8571	26.1081	122	780	2414.72
6	GREEN REGENCY TEACUP AND SAUCER	PICK REGENCY TEACUP AND SAUCER, ROSES REGENCY TEACUP AND SAUCER	0.0213	0.6486	26.1081	122	780	2414.72
7	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0248	0.8485	25.8544	74	401	1295.71
8	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0248	0.7568	25.8544	74	401	1295.71
9	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0248	0.8235	24.4241	82	499	635.25
10	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS	0.0248	0.7368	24.4241	82	499	635.25
11	ROLL WRAP 50'S RED CHRISTMAS	ROLL WRAP 50'S CHRISTMAS	0.0240	0.8710	23.9410	72	610	777.63
12	ROLL WRAP 50'S CHRISTMAS	ROLL WRAP 50'S RED CHRISTMAS	0.0240	0.6585	23.9410	72	610	777.63
13	SET 12 COLOUR PENCILS SPACEBOY	SET 12 COLOUR PENCILS DOLLY GIRL	0.0257	0.7632	23.8911	75	1493	989.65
14	SET 12 COLOUR PENCILS DOLLY GIRL	SET 12 COLOUR PENCILS SPACEBOY	0.0257	0.8056	23.8911	75	1493	989.65
15	CHRISTMAS HANGING STAR WITH BELL	HANGING HEART WITH BELL	0.0204	0.6970	23.8026	72	1891	878.77
16	HANGING HEART WITH BELL	CHRISTMAS HANGING STAR WITH BELL	0.0204	0.6970	23.8026	72	1891	878.77
17	SET OF 6 SNACK LOAF BAKING CASES	SET OF 12 MINI LAOF BAKING CASES	0.0204	0.7419	22.5990	72	496	427.98
18	SET OF 12 MINI LOAF BAKING CASES	SET OF 6 SNACK LOAF BAKING CASES	0.0204	0.7419	22.5990	72	496	427.98
19	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0213	0.8571	21.0000	122	780	2414.72
20	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY TEACUP AND SAUCER	0.0213	0.8571	21.0000	122	780	2414.72

By comparing the product affinity analysis for the previous biweek and the second biweek, the most remarkable observation would be the top spot of the ranking was taken by another rule which included items “REGENCY TEA PLATE GREEN” and “REGENCY TEA PLATE ROSES”. The association rule including the “GLASS APOTHECARY BOTTLE TONIC” and “GLASS APOTHECARY BOTTLE PERFUME” which dominated the top spot of the ranking in the previous biweek exited the ranking. Hence, businesses could take note of the shift in ranking to adjust their inventory level for certain items for such cases. In this biweek, there were getting more rules involving Christmas-related items such as “SET OF 3 WOODEN TREE DECORATIONS”, “SET OF 3 WOODEN STOCKING DECORATION”, “ROLL WRAP 50’S CHRISTMAS”, “ROLL WRAP 50’S RED CHRISTMAS”, “CHRISTMAS HANGING STAR WITH BELL”, and “HANGING HEART WITH BELL” in the ranking. This circumstance further strengthened the trend of Christmas approaching while the dolls-related rules were replaced by the Christmas-related rules. This suggests businesses do not overstock the doll items in the first biweek to minimize the loss. Aside from the regency tea plate series, there were other similar items such as “REGENCY TEACUP AND SAUCER” which came in three different colours, green, roses, and pink in the ranking. They showed up in the ranking with a relatively significant total sales value of 2414.72 which could be considered as the highest sales value in the second biweek. Businesses could create themed displays or promotions featuring all three sets to encourage customers to purchase the complete set and potentially increase sales value. New colour variations for this item could also be considered to offer a better variety in terms of colours for this item and potentially increase sales value. In terms of total sales quantity, the rules which involved “CHRISTMAS HANGING STAR WITH BELL” and “HANGING HEART WITH BELL” had the highest total sales quantity of 1891 units and outperformed other rules in the ranking. However, the total sales value for this rule was relatively low which might be caused by the low unit price of the items, hence, businesses could consider adjusting their pricing strategy such as slightly increasing the price of the items to gain higher revenue.

Table 4.3.5. Top 20 highest frequency itemsets in Biweek 3 (22 October 2011 – 4 November 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0. 0980	131	7863	15235.77
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.0927	124	2309	6757.31
3	WHITE HANGING HEART T-LIGHT HOLDER	0.0733	98	2131	6470.21
4	HOT WATER BOTTLE KEEP CALM	0.0666	89	844	3758.20
5	ASSORTED COLOUR BIRD ORNAMENT	0.0636	85	1668	2879.40
6	POPCORN HOLDER	0.0636	85	3609	2836.67
7	JUMBO BAG PAISLEY PARK	0.0636	85	1418	2738.89
8	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0628	84	1302	436.88
9	GARDENERS KNEELING PAD KEEP CALM	0.0621	83	590	1067.70
10	JUMBO BAG RED RETROSPOT	0.0621	83	1447	2979.36
11	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0606	81	840	2430.08
12	REGENCY CAKESTAND 3 TIER	0.0598	80	331	4199.46
13	LUNCH BAG PAISLEY PARK	0.0591	79	939	1491.83
14	CHOCOLATE HOT WATER BOTTLE	0.0591	79	393	2095.67
15	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0583	78	1635	586.05
16	ALARM CLOCK BAKELIKE GREEN	0.0583	78	732	3019.26
17	SET OF 3 CAKE TINS PANTRY DESIGN	0.0568	76	268	1422.44
18	DOORMAT KEEP CALM AND COME IN	0.0546	73	281	2363.91
19	JUMBO BAG VINTAGE DOILY	0.0539	72	980	2002.65
20	HAND WARMER OWL DESIGN	0.0539	72	759	1577.18

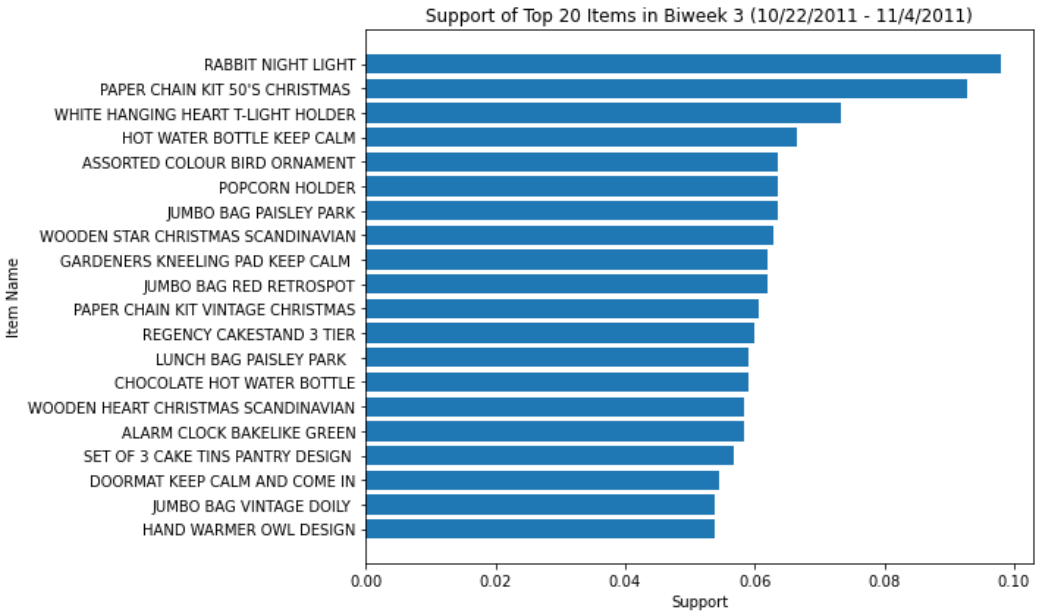


Figure 4.3.3. Visualization of support for top 20 popular items in Biweek 3

As mentioned previously, the "PAPER CHAIN KIT 50'S CHRISTMAS" maintained a high level of support in the third biweek as well. Although the support value slightly declined by 1.5% to 0.0941, however, this item remained one of the top-performing items in the ranking. This insight could further verify the inference of the sustained demand for Christmas-themed items. Businesses could continue to leverage this trend by ensuring adequate stock and promoting this item as a festive decoration to further boost revenue. In terms of ranking, the "RABBIT NIGHT LIGHT" became the top-selling item with a sudden increase in popularity while it had no position in the ranking of the previous biweek. Businesses could make use of this trend to promote this item as a must-have product and potentially bundle it with other complementary items to boost sales further. Furthermore, the "SET OF 3 CAKE TINS PANTRY DESIGN" experienced a significant drop in ranking as it moved from fifth place to seventeenth place in the third biweek. According to this insight, businesses are advised to reassess their pricing strategy or product positioning for this item to regain its popularity among customers. Another notable observation is the "BAKING SET 9 PIECE RETROSPOT" which ranked fourteenth in the previous biweek did not appear in the top ranking of the third biweek. Businesses could explore the reasons behind this decline and adjust their marketing strategies such as introducing new baking sets or promoting alternative baking-related products.

Table 4.3.6. Product affinity analysis for Biweek 3 (22 October 2011 – 4 November 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0209	0.7778	24.7593	147	823	1142.49
2	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.6667	24.7593	147	823	1142.49
3	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.8485	23.6338	147	823	1142.49
4	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	0.0209	0.5833	23.6338	147	823	1142.49
5	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0217	0.8788	23.4988	156	929	1279.87
6	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	0.0217	0.5800	23.4988	156	929	1279.87
7	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.5600	22.6885	147	823	1142.49
8	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0209	0.8485	22.6885	147	823	1142.49
9	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0202	0.7105	22.6184	153	901	1232.77
10	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN HEART DECORATIONS	0.0202	0.6429	22.6184	153	901	1232.77
11	CHARLOTTE BAG PINK POLKADOT	RED RETROSPORT CHARLOTTE BAG	0.0202	0.7500	22.8833	82	849	1229.89
12	RED RETROSPORT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0202	0.6000	22.8833	82	849	1229.89
13	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0247	0.7857	21.8854	97	512	722.28
14	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0247	0.6875	21.8854	97	512	722.28
15	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0202	0.8182	21.8782	153	901	1232.77
16	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	0.0202	0.5400	21.8782	153	901	1232.77
17	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0217	0.7632	21.2571	156	929	1279.87
18	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	0.0217	0.6042	21.2571	156	929	1279.87
19	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0247	0.7857	21.0100	97	553	757.80
20	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS	0.0247	0.7857	21.0100	97	553	757.80

In the third biweek, the top spot of the ranking was taken by the rule involving the items “SET OF 3 WOODEN STOCKING DECORATION”, “SET OF 3 WOODEN SLEIGH DECORATIONS”, and “SET OF 3 WOODEN TREE DECORATIONS” with a support value of 0.0209 and a lift of 24.793. It was noteworthy that the rule involving these three items did not appear in the ranking of the previous biweek, however, it took the first place in the ranking when it showed up. Additionally, there were more rules that included Christmas-related items in the ranking for this biweek while there was only one set of rules which was not related to Christmas in the ranking, which was the rule including items “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG”. Businesses could respond to these trends by adjusting their inventory to meet the changing demands or by offering promotions on these items to capitalize on their popularity. In terms of exits in ranking, the rules related to regency tea plates and teacup sets were all eliminated in this biweek while their position was replaced by the other Christmas-decoration items. Most of the rules in this biweek’s ranking were having their position in the ranking for the first time, hence, their fluctuation in terms of rule’s metrics could not be tracked. To be mentioned that, the Christmas trend started in the previous biweek, which was around the middle of October while the trend was still ongoing in this biweek. For the only not Christmas-related rule, the items “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG” also performed well in this biweek with a support value of 0.0202 and a total sales value of 1229.89. Moreover, the sales quantity for these items was the second highest among the items in the ranking which was 849 units in total. Although the Christmas-related items were popular in this period, however, they did not contribute much to the total sales value and this might be due to the unit price of the items were not that high. Businesses could take this opportunity to adjust their pricing strategy for these items to boost the overall revenue using their popularity. Another notable observation is the rule which had the highest support in the ranking but it positioned the lowest in the ranking. The rule involving “SET OF 3 WOODEN TREE DECORATIONS” and “SET OF 3 WOODEN STOCKING DECORATIONS” had the highest support of 0.0247 which indicates they attracted the most interest from customers but the rule had lower lift.

Table 4.3.7. Top 20 highest frequency itemsets in Biweek 4 (5 November 2011 – 18 November 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0.1577	264	7184	18685.06
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.1123	188	3397	15262.31
3	HOT WATER BOTTLE KEEP CALM	0.0872	146	1072	5639.12
4	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0812	136	1816	7034.32
5	JUMBO BAG PAISLEY PARK	0.0806	135	2020	5630.48
6	JUMBO BAG 50'S CHRISTMAS	0.0806	135	2543	7838.61
7	JUMBO BAG RED RETROSPOT	0.0795	133	3735	7167.05
8	WHITE HANGING HEART T-LIGHT HOLDER	0.0783	131	1900	5647.04
9	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0765	128	2617	863.37
10	LUNCH BAG PAISLEY PARK	0.0723	121	1277	2042.25
11	WOODEN HEART CHRISTMAS SCANDINAVIAN	0.0711	119	2287	750.63
12	JUMBO BAG VINTAGE DOILY	0.0711	119	2463	6609.78
13	BAKING SET 9 PIECE RETROSPOT	0.0711	119	562	2808.46
14	HAND WARMER OWL DESIGN	0.0681	114	1423	3443.78
15	POPCORN HOLDER	0.0657	110	7030	9469.71
16	CHOCOLATE HOT WATER BOTTLE	0.0639	107	1016	4764.90
17	GARDENERS KNEELING PAD KEEP CALM	0.0639	107	1438	3220.10
18	ASSORTED COLOUR BIRD ORNAMENT	0.0603	101	3053	4779.25
19	HOT WATER BOTTLE TEA AND SYMPATHY	0.0591	99	1011	5420.07
20	REGENCY CAKESTAND 3 TIER	0.0591	99	411	5468.13

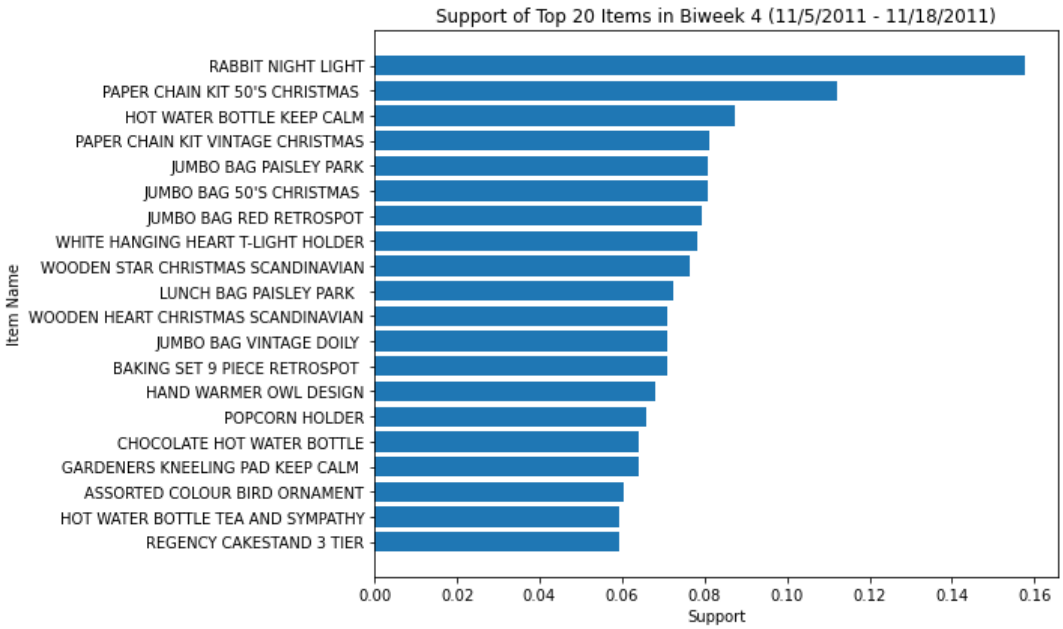


Figure 4.3.4. Visualization of support for top 20 popular items in Biweek 4

In the first two weeks of November, the "RABBIT NIGHT LIGHT" continued its dominance of the top position in the ranking. This item experienced a substantial increase in support value which rose from 0.0980 to 0.1577. The total sales for this item remained high at 7184 with a total sales value of 18685.06. Businesses should take note of the sustained popularity of this item and consider strategies to capitalize on its success such as offering complementary products or promotions. In addition, "PAPER CHAIN KIT 50'S CHRISTMAS" steadily increased its support value from 0.0927 to 0.1123 while the total sales quantity also increased to 3397 and a total sales value of 15262.31. The doubled-up total sales value for this item suggests businesses explore the factors contributing to the increased demand for this item and leverage this trend by optimizing their inventory of seasonal products as well as enhancing marketing efforts to capitalize on festive themes. One notable entry in this biweek is the "HOT WATER BOTTLE TEA AND SYMPATHY" which entered the ranking at the nineteenth position. Despite its lower rank, this item achieved significant sales performance with a total sales quantity of 1011 and a total sales value of 5230.07. This insight suggests a growing demand for comforting and nostalgic products and this surge provides opportunities for businesses to expand their offerings in this category and cater to customer preferences for self-care products, especially during colder seasons.

Table 4.3.8. Product affinity analysis for Biweek 4 (5 November 2011 – 18 November 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0251	0.9545	26.1952	212	1325	1819.16
2	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN HEART DECORATIONS	0.0251	0.6885	26.1952	212	1325	1819.16
3	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN HEART DECORATIONS	0.0251	0.6563	24.4125	212	1325	1819.16
4	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0251	0.9333	24.4125	212	1325	1819.16
5	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0209	0.7292	23.0307	103	591	1886.41
6	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0209	0.6604	23.0307	103	591	1886.41
7	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0209	0.8537	22.3285	194	1040	1436.77
8	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION, SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.5469	22.3285	194	1040	1436.77
9	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0209	0.7955	21.8294	194	1040	1436.77
10	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.5738	21.8294	194	1040	1436.77
11	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS	0.0293	0.8033	21.0108	133	709	975.83
12	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0293	0.7656	21.0108	133	709	975.83
13	CHARLOTTE BAG PINK POLKADOT	RED RETROSPOT CHARLOTTE BAG	0.0221	0.8222	20.8545	113	1578	1975.98
14	RED RETROSPOT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0221	0.5606	20.8545	113	1578	1975.98
15	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0209	0.7143	20.6158	194	1040	1436.77
16	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	0.0209	0.6034	20.6158	194	1040	1436.77
17	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN HEART DECORATIONS	0.0251	0.8571	20.2093	212	1325	1819.16
18	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS, SET OF 3 WOODEN STOCKING DECORATION	0.0251	0.5915	20.2093	212	1325	1819.16
19	SET OF 3 WOODEN SLEIGH DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0263	0.7586	19.8427	129	710	981.92
20	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN SLEIGH DECORATIONS	0.0263	0.6875	19.8427	129	710	981.92

The top spot of the ranking in the fourth biweek was still dominated by Christmas-related rules but the items involved were different from the previous which is the rule involving “SET OF 3 WOODEN TREE DECORATIONS”, “SET OF 3 WOODEN HEART DECORATIONS”, and “SET OF 3 WOODEN STOCKING DECORATIONS”. This rule had the relatively high support of 0.0251 which increased by 24.3% from 0.0202 while the support also increased by 15.7% to 0.0251. The overall sales value in this biweek was higher compared to the previous biweek as the Christmas-related decoration items were getting sold in more quantity. The stable customer preference for these items could allow businesses to promote these items together or offer them as part of a bundled discount to encourage larger purchases. Another notable observation is the rule involving items “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG” remained its position in the ranking without getting eliminated by other Christmas-related rules. The ranking for this rule dropped from twelfth to fourteenth in the ranking, however, the total sales value of these items experienced an increase of 60.7% and this increase could be attributed to the rise of the support value. Businesses should ensure that these items are conveniently located together to enhance the overall customer experience. Furthermore, another interesting insight revealed by the product affinity analysis for the fourth biweek is that the rule involving the “GREEN REGENCY TEACUP AND SAUCER” and “ROSES REGENCY TEACUP AND SAUCER” showed up again in the ranking while it had no places in the top 20 ranking in previous biweek. The resurgence of the teacup sets was attributed to the high lift value of 23.0307 and the stable performance in terms of the total sales value of 1886.41 which was the second-highest total sales value among the rules. Additionally, the rule involving “SET OF 3 WOODEN TREE DECORATIONS” and “SET OF 3 WOODEN STOCKING DECORATION” experienced an increase in support from 0.0245 to 0.0293 while the total sales value was doubled up. Despite there was an increase in support and total sales value, the lift for this rule remained stable at around 21. The consecutively improved performance of the rules involving Christmas-related decoration items suggests businesses continue their efforts on introducing promotions or allocate more space in the store to place a higher volume of the popular items.

Table 4.3.9. Top 20 highest frequency itemsets in Biweek 5 (19 November 2011 – 2 December 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0.1465	236	6687	14025.41
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.1291	208	3456	10836.28
3	HOT WATER BOTTLE KEEP CALM	0.0962	155	877	4620.15
4	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0906	146	1698	5191.32
5	JUMBO BAG 50'S CHRISTMAS	0.0875	141	1459	3081.97
6	POPCORN HOLDER	0.0863	139	6120	5275.14
7	WHITE HANGING HEART T-LIGHT HOLDER	0.0788	127	1727	4713.85
8	HAND WARMER OWL DESIGN	0.0701	113	932	1971.34
9	HOT WATER BOTTLE TEA AND SYMPATHY	0.0695	112	751	4572.69
10	GARDENERS KNEELING PAD KEEP CALM	0.0695	112	1115	2376.31
11	VINTAGE DOILY TRAVEL SEWING KIT	0.0664	107	818	1682.50
12	JUMBO BAG RED RETROSPOT	0.0664	107	1877	3901.56
13	REGENCY CAKESTAND 3 TIER	0.0658	106	459	6192.84
14	60 CAKE CASES VINTAGE CHRISTMAS	0.0639	103	1656	1123.66
15	WOODEN STAR CHRISTMAS SCANDINAVIAN	0.0621	100	2386	864.38
16	TRADITIONAL PICK UP STICKS GAME	0.0621	100	663	885.62
17	CHOCOLATE HOT WATER BOTTLE	0.0615	99	603	3453.58
18	JUMBO BAG PAISLEY APRK	0.0615	99	1082	2214.31
19	CHRISTMAS CRAFT LITTLE FRIENDS	0.0608	98	545	1356.92
20	VINTAGE SNAP CARDS	0.0602	97	1209	1058.07

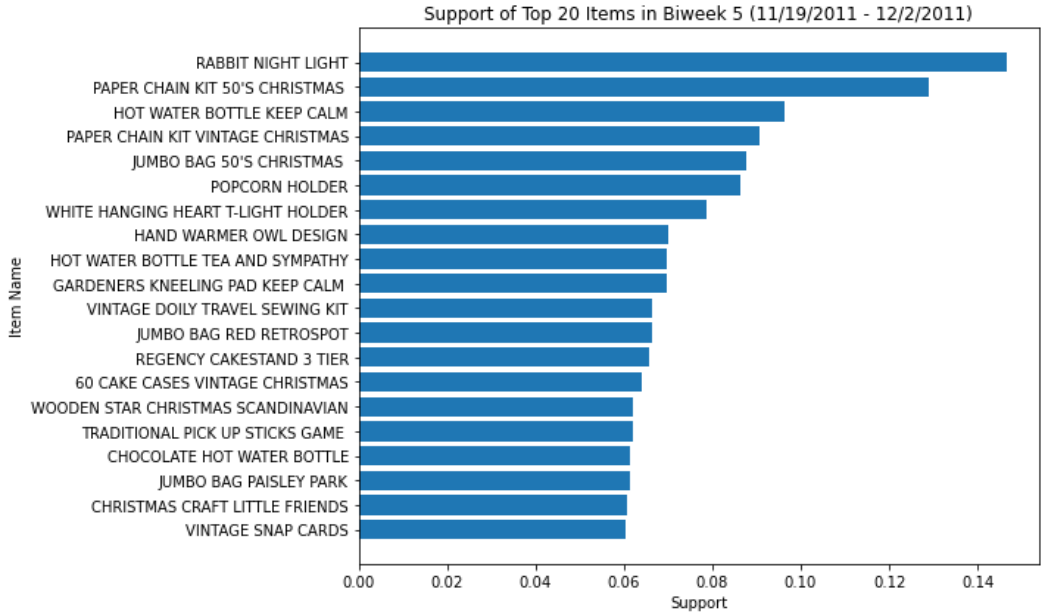


Figure 4.3.5. Visualization of support for top 20 popular items in Biweek 5

The approaching of Christmas and the winter season reveals several interesting trends emerge. The top position of the ranking was still dominated by the “RABBIT NIGHT LIGHT” despite there was a slight decrease in total sales quantity from 7184 to 6687 and total sales value from 18685.06 to 14025.41, the item’s support remained high at 0.14465. The slight decrease in total sales quantity and value indicates a possible saturation point and businesses should consider exploring complementary products or variations to maintain interest and boost sales. Another consistent performer was the “PAPER CHAIN KIT 50’S CHRISTMAS” which retained its second-place ranking across the biweeks. This item outperformed other items with an increase in the total sales quantity, total sales value, and support. The strong seasonal demand suggests businesses consider bundling it with other festive items or offering discounts to leverage its steady popularity during the holiday season. Moreover, the “VINTAGE DOILY TRAVEL SEWING KIT” appeared for the first time in the ranking at the eleventh position with a support value of 0.0664. Businesses could take note of this entry as it might suggest a new or renewed interest in vintage-inspired sewing kits among customers. Conversely, the “ASSORTED COLOUR BIRD ORNAMENT” and “BAKING SET 9 PIECE RETROSPOT” which were ranked eighteenth and thirteenth in the previous biweek exited the ranking in this biweek.

Table 4.3.10. Product affinity analysis for Biweek 5 (19 November 2011 – 2 December 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PINK REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0230	0.8605	27.1806	94	465	1666.31
2	GREEN REGENCY TEACUP AND SAUCER	PINK REGENCY TEACUP AND SAUCER	0.0230	0.7255	27.1806	94	465	1666.31
3	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATION	0.0223	0.6792	24.8696	107	520	739.62
4	SET OF 3 WOODEN TREE DECORATION	SET OF 3 WOODEN STOCKING DECORATION	0.0223	0.8182	24.8696	107	520	739.62
5	ROSES REGENCY TEACUP AND SAUCER	GREEN REGENCY TEACUP AND SAUCER	0.0217	0.7609	24.0345	97	516	1866.12
6	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER	0.0217	0.6863	24.0345	97	516	1866.12
7	CHARLOTTE BAG PINK POLKADOT	RED RETROSPOT CHARLOTTE BAG	0.0211	0.790ass7	23.5891	102	1885	2679.11
8	RED RETROSPOT CHARLOTTE BAG	CHARLOTTE BAG PINK POLKADOT	0.0211	0.6296	23.5891	102	1885	2679.11
9	SET OF 3 WOODEN SLEIGH DECORATION	SET OF 3 WOODEN STOCKING DECORATION	0.0205	0.7174	21.8060	113	571	826.44
10	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN SLEIGH DECORATION	0.0205	0.6226	21.8060	113	571	826.44
11	SPACEBOY LUNCH BOX	DOLLY GIRL LUNCH BOX	0.0230	0.6727	20.4484	119	950	1812.89
12	DOLLY GIRL LUNCH BOX	SPACEBOY LUNCH BOX	0.0230	0.6981	20.4484	119	950	1812.89
13	SET 12 COLOUR PENCILS DOLLY GIRL	SET 12 COLOUR PENCILS SPACEBOY	0.0230	0.7551	20.2745	111	999	792.75
14	SET 12 COLOUR PENCILS SPACEBOY	SET 12 COLOUR PENCILS DOLLY GIRL	0.0230	0.6167	20.2745	111	999	792.75
15	CLASSIC BICYCLE CLIPS	BICYCLE PUNCTURE REPAIR KIT	0.0279	0.7895	20.1880	131	1380	2514.25
16	BICYCLE PUNCTURE REPAIR KIT	CLASSIC BICYCLE CLIPS	0.0279	0.7143	20.1880	131	1380	2514.25
17	ROLL WRAP VINTAGE CHRISTMAS	ROLL WRAP 50'S CHRISTMAS	0.0248	0.6557	17.0386	130	817	1338.27
18	ROLL WRAP 50'S CHRISTMAS	ROLL WRAP VINTAGE CHRISTMAS	0.0248	0.6452	17.0386	130	817	1338.27
19	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN HEART DECORATIONS	0.0205	0.6226	16.1786	131	745	1070.72
20	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0205	0.6226	16.1786	131	745	1070.72

By conducting product affinity analysis for the fifth biweek, it can be observed that the top spot of the ranking was taken by the itemsets “PINK REGENCY TEACUP AND SACUER” and “GREEN REGENCY TEACUP AND SACUER” with a lift value of 27.1806 and support of 0.0230. According to this insight, businesses could collaborate with local cafes or tea houses to feature these tea cups in their establishments for better visibility to the public. The overall product affinity analysis showed that the rules related to Christmas-decoration items were getting lesser in the ranking according to the top 20 rule while there were only 4 sets of rules remaining in this biweek. However, the association rule which involved “SET OF 3 WOODEN STOCKING DECORATION” and “SET OF 3 WOODEN TREE DECORATIONS” remained at a high position in the ranking with a lift of 24.8696. This is an interesting insight to be mentioned as this rule was having 0.0223 which implies that it experienced a substantial decrease of 23.9% for its support. Businesses could take note of this insight and consider decreasing inventory levels for other Christmas-related items while continue stocking up for these two items. As mentioned previously, the rule involving “CHARLOTTE BAG PINK POLKADOT” and “RED RETROSPOT CHARLOTTE BAG” remained its position in the ranking and it even got a higher position from thirteenth to the top 10. The shift in ranking was attributed to the its high lift as well as the total sales value which rose by 35.6% to 2679.11. Furthermore, there were several new entries in the ranking such as the rule involving “BICYCLE PUNTURE REPAIR KIT” and “CLASSIC BICYCLE CLIPS” showed up in the ranking with a lift value of 20.1880 which attributed to the complementary nature between these items. Businesses could create bundled promotions or discounts for bicycle accessories for customers to purchase more similar items together. Moreover, the rule with items “SPACEBOY LUNCH BOX” and “DOLLY GIRL LUNCH BOX” entered the ranking for the first time as well. This rule had a support of 0.0230 and a lift of 20.4484 which were considered as relatively high compared to other rules. Businesses are suggested to create a lunch box collection featuring both designs to market to families or individuals looking for unique and fun lunch accessories. They could also offer a loyalty program where customers collect points for purchasing these lunch boxes which could be used to redeem upcoming lunch boxes with new designs for free.

Table 4.3.11. Top 15 highest frequency itemsets in Biweek 6 (3 December 2011 – 15 December 2011)

	Itemsets	Support	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	RABBIT NIGHT LIGHT	0.1321	93	3380	7584.48
2	PAPER CHAIN KIT 50'S CHRISTMAS	0.1222	86	1545	5328.99
3	HAND WARMER OWL DESIGN	0.0952	67	689	1487.66
4	HOT WATER BOTTLE KEEP CALM	0.0938	66	480	2902.84
5	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.0866	61	666	2074.38
6	GARDENERS KNEELING PAD KEEP CALM	0.0810	57	573	1372.97
7	CHOCOLATE HOT WATER BOTTLE	0.0795	56	360	2367.55
8	BICYCLE PUNCTURE REPAIR KIT	0.0795	56	1109	2265.91
9	POPCORN HOLDER	0.0739	52	4563	4084.27
10	HAND WARMER RED LOVE HEART	0.0724	51	305	654.71
11	HAND WARMER BIRD DESIGN	0.0710	50	668	1375.92
12	JUMBO BAG 50'S CHRISTMAS	0.0710	50	783	1636.89
13	SCOTTIE DOG HOT WATER BOTTLE	0.0682	48	304	1587.65
14	REGENCY CAKESTAND 3 TIER	0.0682	48	388	5170.50
15	GARDENERS KNEELING PAD CUP OF TEA	0.0682	48	380	706.44
16	HAND WARMER SCOTTY DOG DESIGN	0.0668	47	485	1039.52
17	RETROSPOT TEA SET CERAMIC 11 PC	0.0653	46	170	988.22
18	WHITE HANGING HEART T-LIGHT HOLDER	0.0639	45	619	1770.69
19	GIANT 50'S CHRISTMAS CRACKER	0.0639	45	311	422.05
20	HOT WATER BOTTLE TEA AND SYMPATHY	0.0639	45	492	3149.96

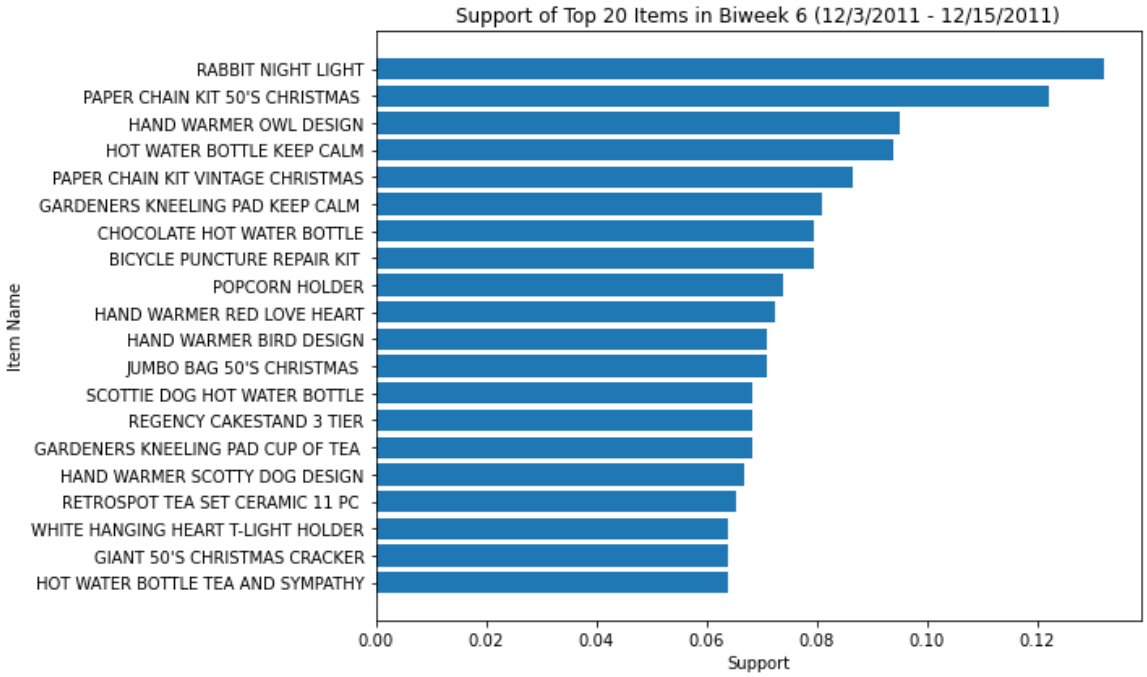


Figure 4.3.6. Visualization of support for top 20 popular items in Biweek 6

In the transition from the previous biweek to the sixth biweek, the popular item “RABBIT NIGHT LIGHT” experienced a significant decrease in total sales quantity and value which signals the potential changes in customer demand or increased competition in the market. The total sales quantity and total sales value for this item declined by 49.51% and 46.03% respectively. This indicates a potential market saturation for this item and businesses should consider diversifying their product range or enhancing marketing efforts to maintain visibility and appeal to customers. Exploring new design variations could be a suggested option to rejuvenate interest in the market. Similarly, the Christmas-themed item, “PAPER CHAIN KIT 50’S CHRISTMAS” saw a substantial decline in total sales quantity from 3456 to 1545 in the sixth biweek which was approximately 64.32% while the total sales value declined by around 50% as well. Despite these drastic decreases, the item’s support only decreased slightly from 0.1291 to 0.1222 which indicates that this item still potential to continue its popularity among customers, but just with lower sales volume. Moreover, seasonal items such as “HOT WATER BOTTLE KEEP CALM” and “HAND WARMER OWL DESIGN” experienced a decline in sales metrics as well. Businesses should precisely plan the stocking level of these seasonal items to avoid overstocking and minimize the loss as much as possible.

Table 4.3.12. Product affinity analysis for Biweek 6 (3 December 2011 – 15 December 2011)

	Antecedent	Consequent	Support	Confidence	Lift	Total Receipt Count	Total Sales Quantity	Total Sales Value
1	PINK DOG BOWL	PINK CAT BOWL	0.0213	0.8824	34.5098	35	174	176.83
2	PINK CAT BOWL	PINK DOG BOWL	0.0213	0.8333	34.5098	35	174	176.83
3	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0213	0.9375	30.0000	72	401	624.99
4	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN TREE DECORATIONS	0.0213	0.9375	30.0000	71	402	599.30
5	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN STOCKING DECORATIONS	0.0213	0.6818	30.0000	72	401	624.99
6	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN STOCKING DECORATIONS	0.0213	0.6818	30.0000	71	402	599.30
7	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	0.0213	0.6522	28.6957	72	401	624.99
8	SET OF 3 WOODEN STOCKING DECORATIONS	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	0.0213	0.6522	28.6957	71	402	599.30
9	SET OF 3 WOODEN HEART DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0213	0.9375	28.6957	72	401	624.99
10	SET OF 3 WOODEN SLEIGH DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS	0.0213	0.9375	28.6957	71	402	599.30
11	SET OF 6 SNACK LOAF BAKING CASES	SET OF 12 MINI LOAF BAKING CASES	0.0227	0.9412	26.5035	48	224	229.93
12	SET OF 12 MINI LOAF BAKING CASES	SET OF 6 SNACK LOAF BAKING CASES	0.0277	0.6400	26.5035	48	224	229.93
13	SET OF 3 WOODEN STOCKING DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN HEART DECORATIONS	0.0213	0.7895	26.4662	72	401	624.99
14	SET OF 3 WOODEN HEART DECORATIONS	SET OF 3 WOODEN STOCKING DECORATIONS, SET OF 3 WOODEN TREE DECORATIONS	0.0213	0.7143	26.4462	72	401	624.99
15	SET OF 3 WOODEN TREE DECORATIONS	SET OF 3 WOODEN STOCKING DECORATION	0.0270	0.8636	26.4348	48	243	363.04
16	SET OF 3 WOODEN STOCKING DECORATION	SET OF 3 WOODEN TREE DECORATIONS	0.0270	0.8261	26.4348	48	243	363.04
17	ROSES REGENCY TEACUP AND SAUCER, SET OF 3 REGENCY CAKE TINS	GREEN REGENCY TEACUP AND SAUCER	0.0213	0.8824	25.8824	86	743	3031.01
18	ROSES REGENCY TEACUP AND SAUCER, SET OF 3 REGENCY CAKE TINS	PINK REGENCY TEACUP AND SAUCER	0.0213	0.8824	25.8824	86	675	2867.73
19	GREEN REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, SET OF 3 REGENCY CAKE TINS	0.0213	0.6250	25.8824	86	743	3031.01
20	PINK REGENCY TEACUP AND SAUCER	ROSES REGENCY TEACUP AND SAUCER, SET OF 3 REGENCY CAKE TINS	0.0213	0.6250	25.8824	86	675	2867.73

In the last biweek, the top spot of the ranking changed again and it was taken by a new rule appearing in the ranking which contained the items “PINK DOG BOWL” and “PINK CAT BOWL”. Although this rule had an extremely high lift and support value, the total sales value for the items was only 176.83 which did not contribute much to the overall revenue in this biweek period. The focus shifts toward pet-related products indicate a preference for matching pet accessories which suggest businesses could take the opportunity to expand the colour variations for this item and promote them as a stylish and coordinated option for pet owners looking to enhance their home decorations. In addition, the rules related to Christmas decoration items experienced a resurgence in this period and most of them returned to the ranking. For instance, the itemsets involving “SET OF 3 WOODEN HEART DECORATIONS”, “SET OF 3 WOODEN STOCKING DECORATION”, and “SET OF 3 WOODEN TREE DECORATIONS” obtained the third position in the ranking with a lift value of 30.000 while the support was same as the rules related to pet accessories which are mentioned previously. Businesses could highlight the seasonal versatility of these wooden decorations such that they are not only limited to holiday decorations but can also be used year-round for various occasions. Another notable entry in the ranking was the rules including “SET OF 6 SNASCK LOAF BAKING CASES” and “SET OF 12 MINI LOAF BAKING CASES” which entered the ranking with a lift of 26.5035 and a support of 0.0227. Similarly, businesses could offer a bundle discount for customers purchasing both types of baking cases which cater to baking enthusiasts looking for a variety of options. In addition, there were certain interesting rules showed up in the ranking for this biweek. For instance, the items “GREEN REGENCY TEACUP AND SAUCER” and “PINK REGENCY TEACUP AND SAUCER” were not longer being purchased with teacup sets in different colours but they were purchased with “SET OF 3 REGENCY CAKE TINS”. As suggestions, businesses leverage the association between tea cups and cake tins to cross-promote with baking accessories such as mixing bowls, measuring spoons, or recipe books. Subsequently, the targeted customers are no longer limited to tea enthusiasts but also home bakers as they could be provided with everything they need to host a successful tea party and bake delectable treats.

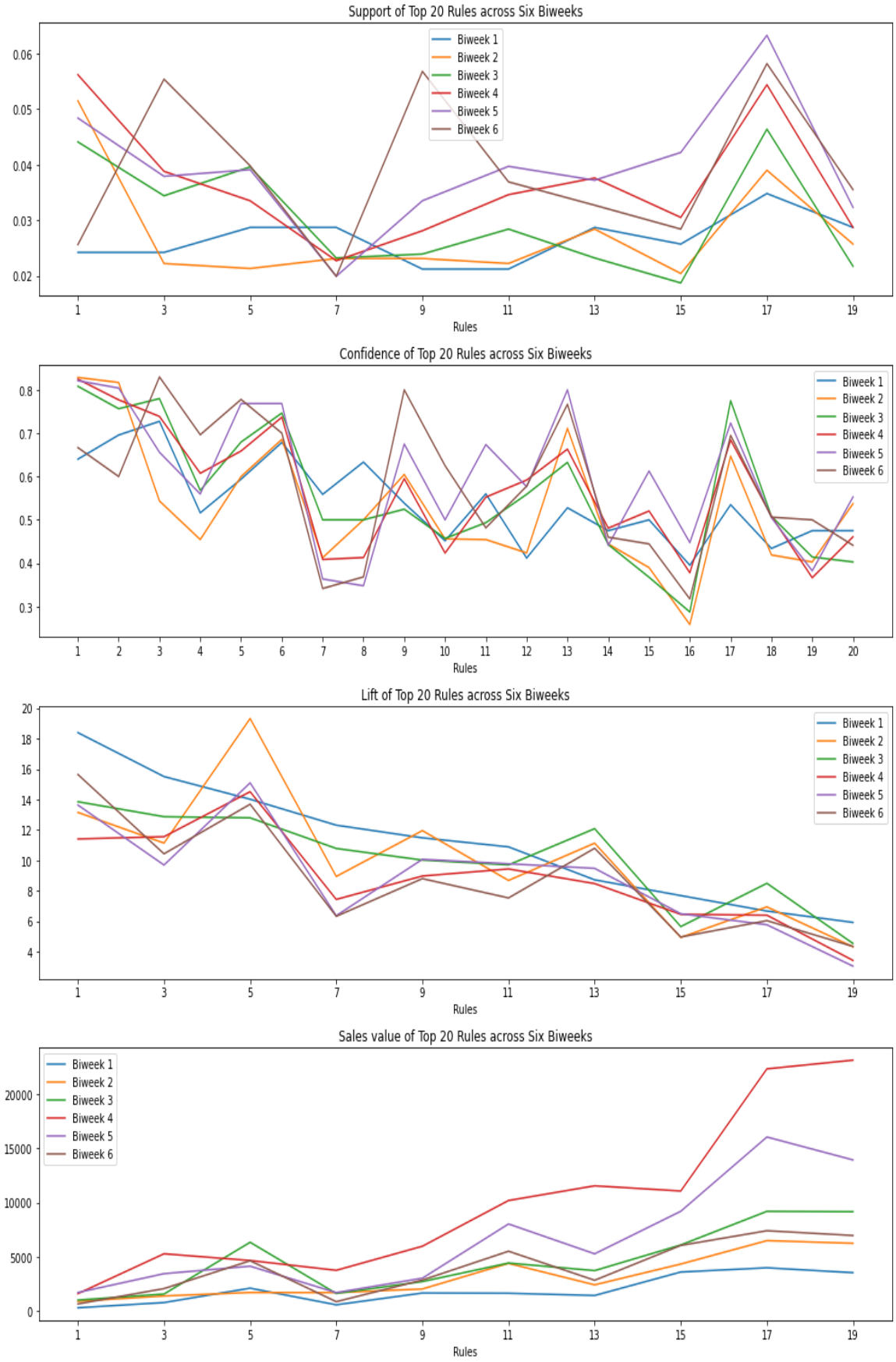


Figure 4.3.7. Metrics fluctuation across six biweeks

According to the fluctuation of metrics visualized by the graphs above, several rules exhibit significant fluctuations in their metrics. First and foremost, Rules 3 and 4 involving the itemsets "GARDENERS KNEELING PAD CUP OF TEA" and "GARDENERS KNEELING PAD KEEP CALM" demonstrated notable fluctuations in support across the six biweeks reaching a peak of 0.0544 in the sixth biweek. This peak suggests a period of increased association between the two items and the steady increase attributed to a surge in demand or a successful marketing campaign. To capitalize on this peak and maintain consistent sales levels, businesses are advised to monitor customer trends and adjust their inventory levels and marketing strategies accordingly for better performance. Conversely, Rules 7 and 8 involving " SET OF 20 VINTAGE CHRISTMAS NAPKINS" and "60 CAKE CASES VINTAGE CHRISTMAS" experienced a steady decrease in support and lift across the six biweeks. The rules reached their peak support of 0.0287 in the first biweek but eventually dropped constantly to 0.0199 in the sixth biweek. This indicates a period of strong association between the two items early on which may have been influenced by seasonal factors or targeted promotions, however, the performance did not last long. Businesses could leverage the early peak performance and sustain the sales momentum by introducing a diverse product range, optimizing marketing strategies, and analyzing customer feedback. Furthermore, Rules 9 and 10 which involved the itemsets "HAND WARMER BIRD DESIGN" and "HAND WARMER OWL DESIGN" showed a notable increase in support over the six biweeks and reached its peak of 0.0567 in the sixth biweek. This consistent upward trend reflects a growing association between the two items which indicates a steady increase in customer interest or demand for these products. This trend could be attributed to seasonal changes as colder weather could increase the demand for hand warmers which eventually lead to high sales of both the bird and owl designs. Similarly, Rules 11 and 12 involving the items "CHOCOLATE HOT WATER BOTTLE" and "HOT WATER BOTTLE TEA AND SYMPATHY" exhibit a constant increase in both support across the biweeks. The rules reached their peak support of 0.0397 in the fifth biweek. Last but not least, most of the rules in the graph contributed the most sales value at the fourth biweek compared to other biweeks. This may be attributed to the beginning of the seasonal trends such as weather changes which made the comfort and warmth provided by these items meet the demand in the market during colder months.

4.4 Comparison between MBA in different periods

This chapter has discussed the product affinity results which obtained by conducting market basket analysis at different time periods such as quarterly, monthly, and biweekly on a dataset containing one year of transaction data. The market basket analysis has highlighted the dynamic nature of customer behavior, the impact of time on purchasing patterns, the influence of seasonal trends, promotional activities, and external factors on product associations and sales patterns. In the context of conducting market basket analysis, the choice of the interval is crucial and should be driven by several key considerations. Quarterly analysis provides several distinct advantages over monthly and biweekly analysis as it could capture long-term trends and patterns in customer behavior. The quarterly analysis offers a broader view of customer preferences and product associations by aggregating transaction data over a three-month period and allows businesses to identify overarching trends that may not be apparent in shorter-term analysis. Moreover, a quarterly MBA provides a relatively more stable and reliable basis for analysis compared to shorter intervals. It helps smooth out short-term fluctuations and noise in data which eventually highlights a more accurate representation of customer behavior over time.

On the other hand, conducting an MBA on a monthly basis provides a balanced view between the granularity of biweekly analysis and the broader perspective of quarterly analysis. The ability to reveal recurring patterns that occur over a few weeks outperforms other periods as the patterns may not be evident in longer-term analysis. These patterns could be attributed to specific events, holidays, or promotional activities that occur on a monthly basis. In addition, monthly MBA also offers a more manageable and actionable level of detail compared to biweekly analysis which might be too frequent for certain businesses to conduct and respond to effectively. The balance between capturing short-term fluctuations and providing a comprehensive overview of customer behavior over a reasonable timeframe allows businesses to optimize their marketing campaigns and inventory management strategies without being overwhelmed by the constant fluctuations seen in biweekly analysis. In contrast, biweekly intervals offer a more granular view of transaction data for businesses to observe short-term fluctuations and trends in customer purchasing behavior. The level of detail is essential to identify immediate and sudden market influences as businesses could track changes in product associations and customer preferences more closely. For instance, if certain product combinations display a sudden increase in support during a specific biweekly period, this could indicate a temporary trend or promotional opportunity that businesses could capitalize on. The

level of insight obtained from a biweekly MBA that may not be apparent in quarterly or monthly analyses is crucial for businesses to stay ahead of changing market dynamics and maintain a competitive edge among the competitors in the market. The earlier adjusting of the marketing strategies according to the market trends provides more opportunities for businesses to capitalize on.

The histogram shown below visualizes the distribution of rules from monthly MBA across different OCVR (Overall Combined Variability Ratio) intervals to provide insightful observations. The majority of the rules fall within the intervals of 1-10% and 10-20%. The concentration of rules in the 1-10% interval shows that a significant number of rules with relatively lower OCVR values. These rules could be considered as less impactful individually contributing to the overall basket behavior. The histogram also highlights a comparatively smaller number of rules falling within the 20-30%. The relatively high OCVR value indicates the rules are very vulnerable to changes and could not be used at any time. This could be attributed to a potential decline in the strength of associations or a reduction in the number of instances where rules exhibit exceptionally high validation rules. t is interesting to note that there are no rules in the intervals beyond 30% which indicates that it is extremely rare for rules to exhibit a high degree of variability across the three-month period.

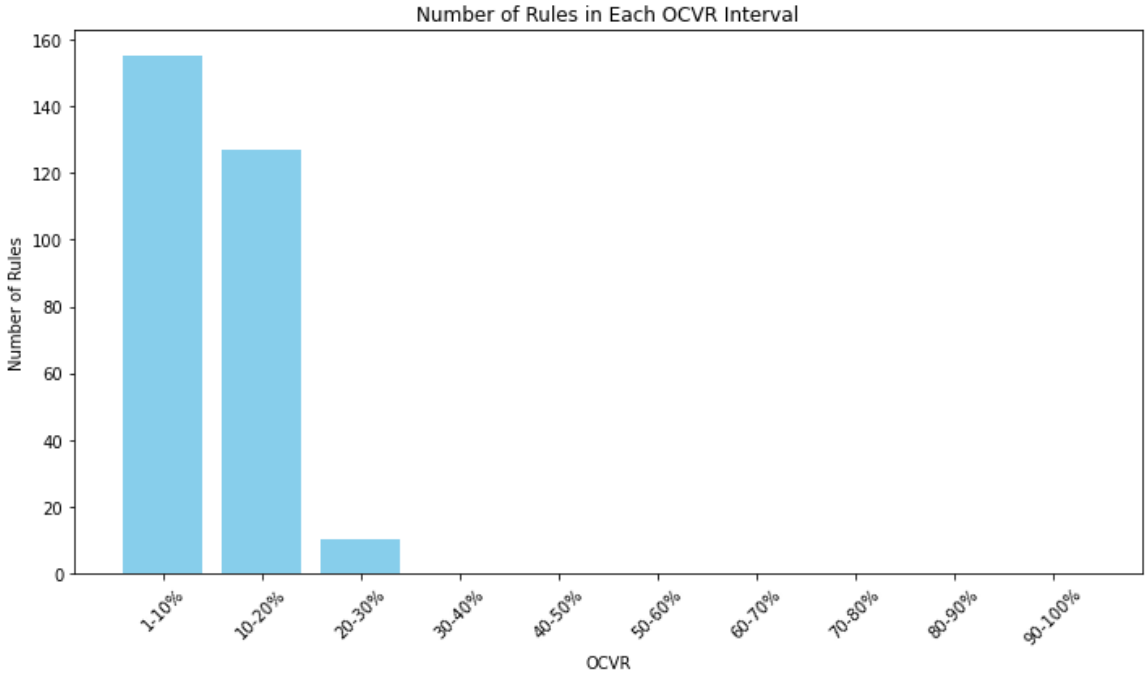


Figure 4.4.1. Histogram of OCVR of monthly rules

Table 4.4.1. Overall variability of top 20 monthly rules

No	Rules	CVL%	CVC%	OCVR%
1	ROSES REGENCY TEACUP AND SAUCER – GREEN REGENCY TEACUP AND SAUCER > PINK REGENCY TEACUP AND SACUER	7.82	3.10	5.46
2	PINK REGENCY TEACUP AND SAUCER > ROSES REGENCY TEACUP AND SAUCER – GREEN REGENCY TEACUP AND SAUCER	7.82	6.20	7.04
3	GREEN REGENCY TEACUP AND SAUCER > PINK REGENCY TEACUP AND SAUCER – ROSES REGENCY TEACUP AND SAUCER	3.39	10.49	6.94
4	PINK REGENCY TEACUP AND SAUCER – ROSES REGENCY TEACUP AND SAUCER > GREEN REGENCY TEACUP AND SAUCER	3.39	1.18	2.28
5	PINK REGENCY TEACUP AND SAUCER > GREEN REGENCY TEACUP AND SAUCER	5.65	1.33	3.49
6	GREEN REGENCY TEACUP AND SAUCER > PINK REGENCY TEACUP AND SAUCER	5.65	5.14	5.40
7	PINK REGENCY TEACUP AND SAUCER – GREEN REGENCY TEACUP AND SAUCER > ROSES REGENCY TEACUP AND SAUCER	3.20	6.92	5.06
8	ROSES REGENCY TEACUP AND SAUCER > PINK REGENCY TEACUP AND SAUCER – GREEN REGENCY TEACUP AND SAUCER	3.20	12.58	7.89
9	SET OF 3 WOODEN TREE DECORATIONS > SET OF 3 WOODEN STOCKING DECORATION – SET OF 3 WOODEN HEART DECORATIONS	12.81	3.94	8.37
10	SET OF 3 WOODEN STOCKING DECORATION – SET OF 3 WOODEN HEART DECORATIONS > SET OF 3 WOODEN TREE DECORATIONS	12.81	6.05	9.43
11	SET OF 3 WOODEN TREE DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS > SET OF 3 WOODEN STOCKING DECORATION	13.97	6.33	10.15
12	SET OF 3 WOODEN STOCKING DECORATION > SET OF 3 WOODEN TREE DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS	13.97	6.87	10.42
13	SET OF 3 WOODEN TREE DECORATIONS > SET OF 3 WOODEN STOCKING DECORATION – SET OF 3 WOODEN SLEIGH DECORATIONS	17.71	5.17	11.44
14	SET OF 3 WOODEN STOCKING DECORATION – SET OF 3 WOODEN SLEIGH DECORATIONS > SET OF 3 WOODEN TREE DECORATIONS	17.71	8.31	13.01
15	PINK REGENCY TEACUP AND SAUCER > ROSES REGENCY TEACUP AND SAUCER	2.69	5.29	3.99
16	ROSES REGENCY TEACUP AND SAUCER > PINK REGENCY TEACUP AND SAUCER	2.69	11.63	7.16
17	SET OF 3 WOODEN SLEIGH DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS > SET OF 3 WOODEN TREE DECORATIONS	18.60	9.43	14.02
18	SET OF 3 WOODEN TREE DECORATIONS > SET OF 3 WOODEN SLEIGH DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS	18.60	7.54	13.07
19	SET OF 3 WOODEN SLEIGH DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS > SET OF 3 WOODEN STOCKING DECORATION	14.50	7.40	10.95
20	SET OF 3 WOODEN STOCKING DECORATION > SET OF 3 WOODEN SLEIGH DECORATIONS – SET OF 3 WOODEN HEART DECORATIONS	14.50	7.27	10.89

According to the table above, Rule 3 which involved the items “GREEN REGENCY TEACUP AND SAUCER”, “PINK REGENCY TEACUP AND SAUCER”, and “ROSES REGENCY TEACUP AND SAUCER” exhibits a relatively low CVL% of 3.39 but the high CVC% of 10.49. The low CVL% reflects the rule has strong stability while the high CVC% suggests that this rule’s confidence varies significantly over time. The OCVR% of 6.94 falls within the 1-10% interval which reflects a moderate overall variability. In contrast, Rule 4 shows stable confidence across the 3-month periods as its low CVC% of 1.18 tells that the confidence in this rule does not vary significantly over time. Similarly, Rule 5 exhibits low CVL%, CVC%, and OCVR% of 5.65, 1.33, and 3.49 respectively which could imply that the popularity of the items related to this rule remains stable over time. In other words, rules with higher CVL% and CVC% may be more sensitive to changes while those with lower CVL% and CVC% are more stable and reliable.

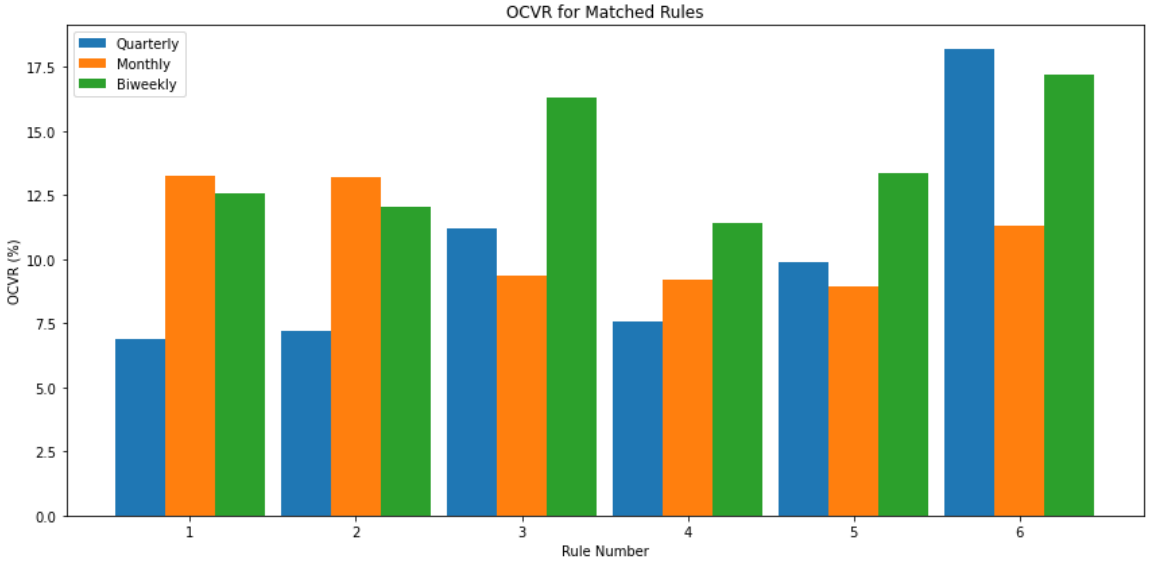


Figure 4.4.2. Overall variability of rules in different periods

The bar graph above illustrates the overall variability of association rules for different periods. The graph clearly shows that most rules exhibit a higher OCVR for the biweekly period compared to the monthly and quarterly periods. The higher variability in the biweekly periods indicates that there is great diversity in customer purchasing patterns within the biweekly periods. Based on these insights, it can be concluded that the biweekly period is a better period to conduct market basket analysis compared to other periods such as quarterly and monthly. The higher OCVR and greater variability in rules’ metrics for the biweekly period show that this period provides more meaningful and actionable insights into customer purchasing behavior.

In conclusion, the choice of interval for MBA should be based on the specific objectives of the analysis and the desired level of detail. If the businesses are mainly focusing on identifying seasonal trends and long-term patterns, quarterly analysis would be more suitable compared to other intervals. Conversely, if the aim is to capture short-term fluctuations and immediate market influences, biweekly or monthly may be more appropriate. According to this dataset, conducting a biweekly MBA would likely be the most beneficial approach for the retail operator by considering their size and resources. By observing the transaction count and the revenue of the operator, a biweekly MBA may not be too expensive to conduct which makes it a suitable choice for a large retail operator with ample resources. Overall, the biweekly MBA allows for a more detailed and timely analysis of transaction data which allows businesses to adjust their marketing strategies and operational decisions quicker for higher effectiveness to drive business growth.

4.5 Objective Reevaluation

The first objective of the project was to conduct market basket analysis in different time periods and this objective was successfully achieved by executing MBA for quarterly, monthly, and biweekly periods through meticulous data processing and analysis. This involved segmenting transaction data into distinct time frames and applying association rule mining techniques to reveal patterns in customer purchasing behavior within each period. An analysis with a focus on identifying frequent itemsets and association rules was conducted to provide valuable insights into the changing market trend. In addition, the second objective which was to analyze the fluctuation of the customer purchasing behaviors with the changes in time was accomplished as well. This was done by visualizing key metrics such as support, lift, confidence, and total sales value for association rules across different time periods. These visualizations provided a comprehensive view of how customer behaviors varied over time and offered insights into potential seasonal trends and purchasing patterns that could inform marketing strategies. Last but not least, the third objective of the project was to provide actionable recommendations for the optimal time period for market basket analysis. This was achieved through a comparative analysis of the overall variability of association rule (OCVR) values across different time periods The analysis successfully identified the biweekly period as the optimal time frame for conducting an MBA compared to other time periods due to its stability and consistency in capturing meaningful insights into customer behavior.

CHAPTER 5 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This project aimed to achieve its objectives of conducting an MBA in different time periods which are quarterly, monthly, and biweekly, analyzing fluctuations in customer purchasing behavior over time, and providing actionable recommendations for the optimal time period to conduct an MBA. Through rigorous data analysis and interpretation, all objectives set forth at the project's outset have been succesfully achieved. The findings of the project provide valuable insights into how customer shopping habits vary across different time periods and offer guidance for businesses seeking to optimize and adjust their marketing strategies and product offerings. The analysis of transaction data categorized into quarterly, monthly, and biweekly reveals distinct patterns in customer purchasing behavior. For instance, certain product associations were more prevalent during specific time periods which could be a good opportunity for cross-selling to be carried out. In terms of actionable recommendations, the findings of the project suggest that the optimal time period for conducting an MBA depends on the specific objectives, resource constraints, and the desired level of responsiveness to market dynamics of the retail business, quarterly analysis provided a broader overview of trends over longer periods which highlights seasonal variations and overall market trends while monthly analysis offered a more detailed perspective compared to quarterly analysis as it able to capture the fluctuations in customer behavior within a single quarter. On the other hand, biweekly analysis provides a granular view of customer purchasing patterns which allows businesses to identify short-term trends and immediate market influences. Larger retail operators with ample resources and a need for real-time insights may conduct an MBA on a biweekly basis as its combination of granularity and frequency enables timely identification of emerging trends and swift adjustments to business strategies. In essence, this project demonstrates the importance of conducting an MBA to gain a comprehensive understanding of customer purchasing behavior. Businesses are advised to conduct MBA in biweekly interval to identify opportunities for potential improvement, optimize their marketing strategies, and enhance the overall customer experience. The insights provided by this project would be valuable for businesses seeking to drive business growth and increase customer satisfaction through data-driven-decision-making.

5.2 Recommendations

According to the findings and methodologies of this project, several recommendations could be implemented upon this project as future work to further enhance the understanding of customer behavior dynamics and optimize retail operations through market basket analysis. First and foremost, including additional data sources could enrich the analysis and provide deeper insights into customer preferences and behaviors. For instance, integrating demographic data and external factors such as weather patterns or economic indicators might be an advantage to gain a more holistic view of customer behavior patterns and their drivers. Moreover, leveraging advanced analytical techniques such as deep learning and natural language processing could reveal the hidden patterns and correlations within the data which eventually provide more accurate predictions and targeted marketing strategies. Secondly, different product categories or customer segments could be used to conduct comparative analyses to gain valuable insights into the varying dynamics of customer behavior. By stratifying the analysis based on factors such as product type, price sensitivity, or customer loyalty, businesses could adjust their marketing strategies and product offerings to better meet the needs and preferences of different customer segments.

In addition, the analysis could be extended to incorporate real-time data streams and implement dynamic MBA frameworks to allow businesses to adapt quickly to changing market conditions and customer trends. Technologies such as data streaming platforms or automated analytics pipelines allow businesses to continuously monitor and analyze transactional data in real time for agile decision-making and rapid adjustments to marketing strategies and inventory management practices. Furthermore, advanced analytics techniques such as network sequence or sequence mining could uncover complex patterns and relationships within the transactional data which provide more sophisticated insights into customer behavior and preferences. By identifying co-occurrence patterns, cross-selling opportunities, and sequential purchasing behaviors, businesses may tailor their marketing campaigns and develop personalized recommendations to enhance the customer experience and drive sales.

REFERENCES

[1] S. Chaudhary, “Understanding market basket analysis in data mining,” Understanding Market Basket Analysis in Data Mining, https://www.turing.com/kb/market-basket-analysis (accessed Feb. 28, 2023).

[2] “Association rule mining: An overview and its applications,” upGrad blog, https://www.upgrad.com/blog/association-rule-mining-an-overview-and-its-applications/ (accessed Feb. 23, 2023).

[3] T. Kristiana et al., “Association rule implementation using algorithm apriori to analize fishing pattern in Indonesia,” *Journal of Physics: Conference Series*, vol. 1641, no. 1, p. 012072, Nov. 2020. doi:10.1088/1742-6596/1641/1/012072.

[4] R. Ruswati, A. I. Gufroni, and R. Rianto, “Associative analysis data mining pattern against traffic accidents using apriori algorithm,” *Scientific Journal of Informatics*, vol. 5, no. 2, pp. 91–104, Nov. 2018. doi:10.15294/sji.v5i2.16199.

[5] K. Cios, R. Swiniarski, W. Pedrycz, and L. Kurgan, *Data Mining*, 2007. doi:10.1007/978-0-387-36795-8.

[6] A. N. Sagin and B. Ayvaz, “Determination of association rules with Market Basket Analysis: Application in the retail sector,” *Southeast Europe Journal of Soft Computing*, vol. 7, no. 1, May 2018. doi:10.21533/scjournal.v7i1.149.

[7] R. Agrawal and R. Srikant, Fast algorithms for Mining Association rules, https://www.vldb.org/conf/1994/P487.PDF (accessed Apr. 24, 2024).

[8] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Waltham, Ma: Morgan Kaufmann, 2012.

[9] W. P. Nurmayanti et al., “Market basket analysis with apriori algorithm and frequent pattern growth (FP-growth) on outdoor product sales data,” *International Journal of Educational Research & Social Sciences*, vol. 2, no. 1, pp. 132–139, Apr. 2021. doi:10.51601/ijersc.v2i1.45.

[10] A. Khalil Jumaa, S. T. F. Al-Janabi, and N. Abedlqader Ali, “Hiding sensitive association rules over privacy preserving distributed data mining,” *Kirkuk University Journal-Scientific Studies*, vol. 9, no. 1, pp. 59–72, Jun. 2014. doi:10.32894/kujss.2014.89609

[11] J. K, “The FP growth algorithm,” Medium, https://towardsdatascience.com/the-fp-growth-algorithm-1ffa20e839b8 (accessed Mar. 6, 2023).

[12] Y. Verma, “Apriori vs FP-growth in Market Basket Analysis - A Comparative Guide,” Analytics India Magazine, https://analyticsindiamag.com/apriori-vs-fp-growth-in-market-basket-analysis-a-comparative-guide (accessed Mar. 6, 2023).

[13] A. A. Aldino, E. D. Pratiwi, Setiawansyah, S. Sintaro, and A. Dwi Putra, “Comparison of market basket analysis to determine consumer purchasing patterns using FP-growth and APRIORI ALGORITHM,” *2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, Oct. 2021. doi:10.1109/icomitee53461.2021.9650317.

[14] M. Hossain, A. H. Sattar, and M. K. Paul, “Market basket analysis using Apriori and FP growth algorithm,” *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, Dec. 2019. doi:10.1109/iccit48885.2019.9038197.

[15] Ansari and S. Zafar, *Market basket analysis : trend analysis of association rules in different time periods*, 2019..

[16] V. Papavasileiou and A. Tsadiras, “*Time Variations of Association rules in Market Basket Analysis,” IFIP Advances in Information and Communication Technology*, pp. 36–44, 2011. doi:10.1007/978-3-642-23960-1_5.

[17] Alfiqra and A. U. Khasanah, “Implementation of market basket analysis based on overall variability of association rule (OCVR) on product marketing strategy,” *IOP Conference Series: Materials Science and Engineering*, vol. 722, no. 1, p. 012068, Jan. 2020. doi:10.1088/1757-899x/722/1/012068.

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Jan 2024	**Study week no.:** 2
Student Name & ID: Teng Wen Foong, 20ACB01778	
Supervisor: Dr Kh'ng Xin Yi	
Project Title: Market Basket Analysis: Analysis of Association Rule in Different Time Periods	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Successfully conducted market basket analysis for monthly and biweekly period as well as the analysis of the result obtained from monthly MBA.

2. WORK TO BE DONE

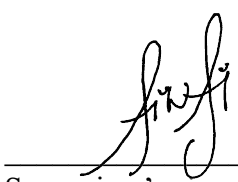
To complete the analysis of the result obtained from biweekly MBA.

3. PROBLEMS ENCOUNTERED

No.

4. SELF EVALUATION OF THE PROGRESS

The progress made has been satisfactory.



Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Jan 2024	**Study week no.:** 4
Student Name & ID: Teng Wen Foong, 20ACB01778	
Supervisor: Dr Kh'ng Xin Yi	
Project Title: Market Basket Analysis: Analysis of Association Rule in Different Time Periods	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Completed the analysis of the result obtained from the biweekly MBA.

2. WORK TO BE DONE

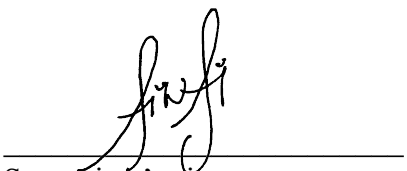
Conduct MBA for one additional period which is quarterly period.

3. PROBLEMS ENCOUNTERED

No.

4. SELF EVALUATION OF THE PROGRESS

The progress is on the track, however, there is still work to be done which require more additional effort and attention to detail.



Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Jan 2024	**Study week no.:** 6
Student Name & ID: Teng Wen Foong, 20ACB01778	
Supervisor: Dr Kh'ng Xin Yi	
Project Title: Market Basket Analysis: Analysis of Association Rule in Different Time Periods	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Successfully conducted quarterly MBA and obtained the result.

2. WORK TO BE DONE

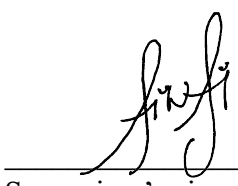
To carry out analysis of the result obtained from the quarterly MBA.

3. PROBLEMS ENCOUNTERED

No.

4. SELF EVALUATION OF THE PROGRESS

The completion of the quarterly MBA marks a significant milestone in the project. The progress made so far is satisfactory.



Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Jan 2024	**Study week no.:** 9
Student Name & ID: Teng Wen Foong, 20ACB01778	
Supervisor: Dr Kh'ng Xin Yi	
Project Title: Market Basket Analysis: Analysis of Association Rule in Different Time Periods	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

Completed the analysis of the result obtained from quarterly MBA.

2. WORK TO BE DONE

To compare the difference between MBA in different period and suggest the optimal period to conduct MBA.

3. PROBLEMS ENCOUNTERED

No.

4. SELF EVALUATION OF THE PROGRESS

The milestone of completing the analysis of the result obtained from quarterly MBA indicates the completion of the project is around the corner. I am dedicated to ensure that the remaining tasks are completed efficiently.

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Jan 2024	**Study week no.:** 12
Student Name & ID: Teng Wen Foong, 20ACB01778	
Supervisor: Dr Kh'ng Xin Yi	
Project Title: Market Basket Analysis: Analysis of Association Rule in Different Time Periods	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

I have completed the comparison between MBA in different period and the optimal period to conduct MBA is figured out.

2. WORK TO BE DONE

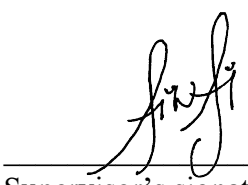
No.

3. PROBLEMS ENCOUNTERED

No.

4. SELF EVALUATION OF THE PROGRESS

The project was completed within the set timeline and met all the objectives outlined at the beginning. I consider the project a success and believe it has contributed positively to my skills and knowledge in the field of market basket analysis.



Supervisor's signature

Student's signature

POSTER

UTAR
UNIVERSITI TUNKU ABDUL RAHMAN

Wholly owned by UTAR Education Foundation
(Co. No. 578227-M)
DU012(A)

MARKET BASKET ANALYSIS: ANALYSIS OF ASSOCIATION RULES IN DIFFERENT TIME PERIODS

OVERVIEW

Dive into the world of consumer behavior and temporal dynamics! Join us on an exploratory journey as we employ cutting-edge market basket analysis techniques to decipher the prime time periods for efficient association rule mining

WHY IT MATTERS?

In an ever-evolving marketplace, understanding when purchasing behaviors peak is gold. This project empowers businesses to allocate resources strategically, tailoring campaigns, and promotions to capitalize on the most opportune moments

METHODOLOGY

1. Conduct MBA in different time periods (quarterly, monthly, biweekly).
2. Analyze and compare difference between MBA in different time periods.
3. Determine the optimal time period to conduct an MBA.

THE BEST PERIOD - BIWEEKLY! WHY NOT OTHER PERIODS?

Biweekly MBA hits the sweet spot between granularity and manageability, offering frequent insights without overwhelming detail. It's like catching the perfect wave of consumer trends, riding the wave of data without getting lost in the sea of daily fluctuations .

PLAGIARISM CHECK RESULT

FYP2

ORIGINALITY REPORT

16%	12%	5%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	rstudio-pubs-static.s3.amazonaws.com Internet Source	3%
2	Submitted to The University of Wolverhampton Student Paper	2%
3	Submitted to University of North Texas Student Paper	2%
4	Submitted to University of Central Florida Student Paper	2%
5	Submitted to CSU, Long Beach Student Paper	1%
6	Submitted to S.P. Jain Institute of Management and Research, Mumbai Student Paper	<1%
7	hdl.handle.net Internet Source	<1%
8	fastercapital.com Internet Source	<1%
9	Submitted to University of Salford	

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 0	Effective Date: 01/10/2013	Page No.: 1of 1

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Full Name(s) of Candidate(s)	Teng Wen Foong
ID Number(s)	20ACB01778
Programme / Course	Bachelor of Computer Science (Honours)
Title of Final Year Project	Market Basket Analysis: Analysis of Association Rules in Different Time Periods

Similarity	**Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)**
Overall similarity index: 16 % **Similarity by source** Internet Sources: 12 % Publications: 5 % Student Papers: 13 %	
Number of individual sources listed of more than 3% similarity: 0	
Parameters of originality required and limits approved by UTAR are as Follows: **(i) Overall similarity index is 20% and below, and** **(ii) Matching of individual sources listed must be less than 3% each, and** **(iii) Matching texts in continuous block must not exceed 8 words** ***Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.***	

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Signature of Supervisor

Name: Dr Kh'ng Xin Yi

Date: 25/4/2024

Signature of Co-Supervisor

Name: _____

Date: _____

UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY
(KAMPAR CAMPUS)

CHECKLIST FOR FYP2 THESIS SUBMISSION

Student Id	20ACB01778
Student Name	Teng Wen Foong
Supervisor Name	Dr Kh'ng Xin Yi

TICK (√)	DOCUMENT ITEMS Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item.
√	Title Page
√	Signed Report Status Declaration Form
√	Signed FYP Thesis Submission Form
√	Signed form of the Declaration of Originality
√	Acknowledgement
√	Abstract
√	Table of Contents
√	List of Figures (if applicable)
√	List of Tables (if applicable)
	List of Symbols (if applicable)
	List of Abbreviations (if applicable)
√	Chapters / Content
√	Bibliography (or References)
√	All references in bibliography are cited in the thesis, especially in the chapter of literature review
	Appendices (if applicable)
√	Weekly Log
√	Poster
√	Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005)
√	I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report.

*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

_____

(Signature of Student)
Date: 25/4/2024