IMPLEMENTATION DATA MINING ON PRODUCT SALES IN PT. FOCUS GAYA GRAHA USING ASSOCIATION RULE METHOD

Aprisal Budiana

Teknik Informatika - Universitas Komputer Indonesia
Jl. Dipatiukur 112-114 Bandung
aprisalbudiana@gmail.com

ABSTRACT

PT. Focus Gaya Graha is a private company specializied in furniture industry.Today the type of the production from PT.Focus Gaya Graha is divided to 3 type,that is wood furniture, panel furniture, and photo frame. Each type of production, produce many type, size, and vary colour. The production system is based on selling history that selling many products that are frequently sold,and this lead to stacking the production from mistake of specifying what product must be produce.

In Data mining,there is some method that named association rule. The purpose of association rule is showing the associative value each products bought by customer so that seen some pattern that forming products that often sold at the same time on one transaction. By knowing the products that often sold at the same time on one transaction,can be made a decision for specifiying what product that effective when produced.

From the test results it can be concluded that the method of data mining association rule can help PT. Focus Gaya Graha to determine the purchasing patterns of consumers, so it can help companies determine the products to be manufactured.

Keyword : Data mining, association rule.

1. INTRODUCTION

1.1. Background

PT. Focus Style Graha is a private company engaged in the furniture industry. At present, the type of production in the PT. Focus Style Graha is divided into three, namely wood furniture, panel furniture and photo frames. Each type of production that produces a variety of products with different types, sizes and colors are different. Product-product produced by PT. Focus Style Graha has been sold to the overseas, it certainly generate a lot of sales transaction data. At present, the use of the sales transaction data is limited to making the report for the company and then archived, not used for a basic consideration in determining the products to be manufactured. S istem production runs in PT. Focus Style Graha is based on sales orders and history of the many products that are often sold. The production system is based on sales history many products are often sold often results in the buildup of production, resulting from errors in determining which products should be produced.

Data mining is a series of processes to explore the added value of a data set of knowledge that had been unknown to manually [1]. In the scientific field of data mining, there is a method called association rule. This method is often also called the market Basketball analysis since the beginning derived from the study of the sales transaction database [2]. Association rule aims to show associative value between the products purchased by the customer so that terlihatlah a pattern in the form of what products are often purchased together in a sales transaction. By knowing what products are often purchased together, may be made a basis for a decision to determine which products are effective when produced.

Based on the description above, it is necessary to implement this method of software association rule as a tool to determine the associative value between the products. Hopefully with this software can help the company to determine the purchasing patterns of consumers, which is where the information can be utilized by the company as a basis for consideration in determining the products to be manufactured.

1.2. Purpose and Objective

1.2.1. Purpose

The intent of this research is to implement data mining on product sales at PT. Focus Style Graha using the association rule.

1.2.2. Objective

The expected goal is to help PT. Focus Style Graha to determine the purchasing patterns of consumers, in order to facilitate in determining what products will be produced.

1.3. Data Mining

Data mining is a term used to describe the knowledge discovery in databases [9].

According Pramudiono [1], "Data mining is a series of processes to explore the added value of a data set of knowledge that has not known it manually. "

According to Larose [4] remarkable progress continues in the field of data mining is driven by several factors, among others:

1. The rapid growth in the data set.
2. Data storage in the data warehouse, so that all companies have access to a reliable database.
3. An increase in data access via the web and intranet navigation.
4. Business competencies pressure to increase market share in the economic globalization.
5. Development of software technology for data mining (availability of technology).
6. Great developments in computing capability and capacity of storage media.

1.4. Cross-Industry Standard Process for Data Mining (CRISP-DM)

CRISP-DM was developed in 1996 by analysts from several industries such as Daimler Chrysler, SPSS and NCR. CRISP-DM provides the standard data processing mining as a general problem-solving strategies of the business or research units.

In CRISP-DM, a project of data mining has a life cycle that is divided into six phases. Overall the phase sequence that is adaptive. The next phase in the sequence depends on the output of the previous phase. Antarfase relationship illustrated by arrows. For example, if the process is in the modeling phase. Based on the behavior and characteristics of the model, the process may have to go back to the data phase preparation for further improvement of the data, or move forward to the evaluation phase. Six phases in the CRISP-DM method can be seen in Figure 1:



Figure 1. CRISP DM

The phases of CRISP-DM [4] :

1. Business Understanding
   Business Understanding or domain understanding (research). In this phase, it takes an understanding of the substance of the activities of data mining will be done, needs from a business perspective. Its activities include: setting goals or business goals, understand the business situation, determine the destination of data mining and planning strategies and study schedules.
2. Data Understanding
   Data Understanding or understanding the data is the initial phase of collecting data, studying the data to get to know the data that will be used. This phase try to identify issues related to data quality, detect an interesting subset of the data to make the initial hypothesis.
3. Data Preparation
   Data preparation or preparation data. This phase is often referred to as a labor-intensive phase. Activities undertaken include selecting tables and fields that will be transformed into a new database for the data material mining (raw data sets).
4. Modelling
   Modeling is a decisive phase technique of data mining is used, determine the tools data mining, data techniques mining, algorithms of data mining, determining the optimal parameter value.
5. Evaluation
   Evaluation is the phase of the interpretation of the results of data mining shown in the modeling process in the previous phase. Evaluation is done in depth with the aim of adjusting the model obtained to suit the objectives to be achieved in the first phase.

6. Deployment
   Deployment or deployment phase is the preparation of reports or presentations of the knowledge gained from the evaluation of the data mining process.

1.5. Association Rule Method

Association rule is the technique of data mining techniques to look for relationships between items within a specified dataset. Examples of associative rules we can take from a sales transaction at a store, we can determine how likely a consumer purchases an item along with other items. This method is often also called the market Basketball analysis since the beginning derived from the study of a database transaction.

If we take the example of a purchase of goods in a store obtained form association → rule bread jam. Which means that consumers who buy bread there is a possibility that the consumer will also buy jam.

Association rule has two stages of processing, ie [11]:

1. Looking for the most common combination of an itemset.
2. Defines the Condition and Result (for conditional association rule).

In determining an association rule, there is an interestingness measure (measure of confidence) is obtained from the processing of data with a particular calculation. There are generally two sizes, namely:

1. Support: a measure that shows how much the level of dominance an item / itemset of the overall transaction.This measure determines whether an item / itemset feasible to look for his confidence (eg, of the total transactions there are, how big the dominance of an item indicates that the item A and B items purchased together).
2. Confidence: a measure that shows the relationship between the two items is conditional (eg, calculating the probability of how often the items purchased by the customer and if the customer is buying an item A).

The two measures will be useful in determining the strength of a pattern by comparing the pattern with a minimum value of both parameters specified by the user. If a pattern of two values meet minimum pre-set parameters, then the pattern can be termed as interesting rule or strong rule.

Association basic methodology of analysis is divided into two stages [9] :

1. Analysis of high frequency pattern
   This stage looking for a combination of items that meet the minimum requirements of the value of the support in the database. Support an item value is obtained by the following formula.

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \quad (1)$$

Meanwhile, the support of 2 items obtained from the following formula.

$$Support\ (A,B) = P(A \cap B) \quad (2)$$

$$Support\ (A,B) = \frac{\Sigma Transaksi\ mengandung\ A\ dan\ B}{\Sigma Transaksi} \quad (3)$$

2. Formation of Association Rules

After all high frequency pattern is found, then search for the associative rules that meet the minimum requirements for confidence by calculating confidence associative rule A -> B of the support of high-frequency pattern A and B, using the formula.

$$Confidence = P(B \mid A) = \frac{\Sigma\ Transaksi\ mengandung\ A\ dan\ B}{\Sigma\ Transaksi\ mengandung\ A} x100\% \quad (4)$$

1.6. Equivalence Class Transformation (Eclat) Algorithm [12]

Both Apriori and FP -growth both frequent method mining of collection transactions using the form TID- itemset, where TID is the transaction ID and itemset is a collection of items purchased on the transaction TID. This is known as horizontal data format. In addition, data can be represented in the form of item-TID_set, where an item is an item name and transaction ID TID_set is a collection that contains the item. This is known as vertical data format. Eclat algorithm is an algorithm that uses a vertical data format for representing data.

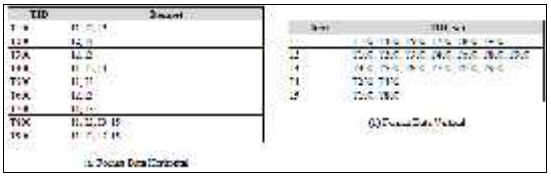


Figure 2. Horizontal and Vertical Data Format

.

Here is an illustration of the process of mining frequent itemset using Eclat algorithm:

1. Represent data into a data format vertical.
2. Determine the minimum value support.
3. Determine the minimum value confidence.
4. Search Data candidate k-itemset ($C_{k)}$ and then calculate each support its value.
5. Remove itemset support its value is less than the minimum value support.
6. After getting the data that meets the minimum value support (frequent), do a combination at the same TID_set itemset based on the data thus creating a k-itemset.

7. Repeat step 4 through step 6 to not generate new itemset.
8. Find asosiatifnya rule by determining the value of his confidence,
9. Remove data that his confidence value is less than the minimum value confidence.

Advantages possessed by Eclat algorithm used in this study is it does not require the calculation of the value of support, because the value of the support of itemset has been represented by TID_set of the itemset. The main job of the Eclat is to combine items based TID_set her, so the size of TID_set is one of the main factors that affect the amount of time and memory required. The larger the TID_set the greater the time and memory required**.**

2. CONTENTS OF RESEARCH

2.1. Analysis of Data Sources

The data used is furniture sales transaction reports in January 2015 to February 2015. The sales transaction report data used furniture is a combination of product data and sales transaction data. Furniture sales transaction reporting data is used the Excel format (* .xls).

Table 1. Samples furniture sales transaction reports

Invoice No.	Date	Item	Qty.	Customer
INVPR2015107801	03/01/2015	LPP.32 - Walnut	2	Galleria Furniture
INVPR2015107801	03/01/2015	LPP.22.M - Italian Walnut	2	Galleria Furniture
INVPR2015107801	03/01/2015	LPP.22 - Beige	1	Galleria Furniture
INVPR2015107802	03/01/2015	LPC.22.R - Beige	2	Meubel Berkah
INVPR2015107802	03/01/2015	MR03 - Italian Walnut	2	Meubel Berkah
INVPR2015107803	03/01/2015	MB.0606 - Beige	2	Citra Jaya Furniture
INVPR2015107803	03/01/2015	RTV.0806 - Beige	2	Citra Jaya Furniture
INVPR2015107804	05/01/2015	BLK.0402.F - Italian Walnut	2	Bintang Jaya Meubel
INVPR2015107804	05/01/2015	BLK.0402 - Beige	2	Bintang Jaya Meubel
INVPR2015107805	05/01/2015	LPS.22 - Walnut	3	Toko Mebel Jadi Jaya
INVPR2015107806	05/01/2015	LHS.0703 - Walnut	2	Toko Mebel Sumber Rezeki
INVPR2015107806	05/01/2015	LHS.0701 - Beige	2	Toko Mebel Sumber Rezeki
INVPR2015107806	05/01/2015	LPC.22.L - Italian Walnut	2	Toko Mebel Sumber Rezeki
...	...	...	...	...

2.2. Data Preparation Analysis

The steps of preparation of data in this study are as follows :

1. Selection Data
   Phase selection of data has the task of covering the election tables and fields. Table selected is a table reporting sales transaction. Field selected is field Invoice No. and items. Field was chosen because of the field can be viewed any products that are frequently purchased consumer simultaneously. Election results data can be seen in Table 2.

Table 2. Result of Selection Data

Invoice No.	Item
INVPR2015107801	LPP.32 - Walnut
INVPR2015107801	LPP.22.M - Italian Walnut
INVPR2015107801	LPP.22 - Beige
INVPR2015107802	LPC.22.R - Beige
INVPR2015107802	MR03 - Italian Walnut
INVPR2015107803	MB.0606 - Beige
INVPR2015107803	RTV.0806 - Beige
INVPR2015107804	BLK.0402.F - Italian Walnut
INVPR2015107804	BLK.0402 - Beige
INVPR2015107805	LPS.22 - Walnut
INVPR2015107806	LHS.0703 - Walnut
INVPR2015107806	LHS.0701 - Beige
INVPR2015107806	LPC.22.L - Italian Walnut
...	...

2. Cleanup Data
   This process is carried out on the data cleansing process of the data that is irrelevant. Criteria relevant data is data that has only one single item in a transaction, because it can be ascertained that the data will not have a linkage with other data. Data cleaning results can be seen in Table 3.

Table 3. Result of Cleanup Data

Invoice No.	Item
INVPR2015107801	LPP.32 - Walnut
INVPR2015107801	LPP.22.M - Italian Walnut
INVPR2015107801	LPP.22 - Beige
INVPR2015107802	LPC.22.R - Beige
INVPR2015107802	MR03 - Italian Walnut
INVPR2015107803	MB.0606 - Beige
INVPR2015107803	RTV.0806 - Beige
INVPR2015107804	BLK.0402.F - Italian Walnut
INVPR2015107804	BLK.0402 - Beige
INVPR2015107806	LHS.0703 - Walnut
INVPR2015107806	LHS.0701 - Beige
INVPR2015107806	LPC.22.L - Italian Walnut
...	...

3. Development Data
   At this step will be the construction of the data, the data that will be built based on the explanation at the stage of understanding the data. The data that will be built is presented in the tables the following.

a. Table penjualan
   Penjualan table is a table that stores data about the sales transaction, the table structure can be seen in table 4.

Tabel 4. Structure of Table penjualan

Field	Tipe Data	Ukuran	Definisi
Invoice No.	Varchar	25	No transaksi pada nota penjualan.
Customer	Varchar	30	Konsumen yang membeli produk.

b. Table detailpenjualan
Detailpenjualan table is a table that stores data about the sales transaction, the table structure can be seen in table 5.

Tabel 5. Structure of Table detailpenjualan

Field	Tipe Data	Ukuran	Definisi
Invoice No.	Varchar	25	No transaksi pada nota penjualan.
Date	Date	-	Tanggal transaksi penjualan.
Item	Varchar	30	Produk yang dibeli oleh konsumen.
Qty.	Int	4	Jumlah produk yang dibeli oleh konsumen.

c. Table preprocessing
Product table is a table that contains information about the product, the table structure can be seen in Table 6.

Tabel 6. Structure of Table preprocessing

Field	Tipe Data	Ukuran	Definisi
Invoice No.	Varchar	25	No transaksi pada nota penjualan.
Item	Varchar	30	Produk yang dibeli oleh konsumen.

4. Format Data
Phase data format is the final stage of preparation of data before starting the modeling phase. This stage has the final form of the data format to be used in the mining process as can be seen in Table 7.

Table 7. The data format for Process Mining

Invoice No.	Item
INVPR2015107801	LPP.32 - Walnut
INVPR2015107801	LPP.22.M - Italian Walnut
INVPR2015107801	LPP.22 - Beige
INVPR2015107802	LPC.22.R - Beige
...	...

2.3. Analysis Modelling

This stage consists of two phases, namely :

a. Modeling Techniques
Modeling techniques are used in accordance with the original purpose of which is to determine the products to be manufactured. The model used is the algorithm Equivalence Class Transformation (Eclat).

b. Testing Analysis Model
Cases that will be tested using Eclat algorithm is as follows:

1. The data used is furniture sales transaction report period January to February 2015 that has been formatted in accordance with the needs of mining process.
2. Represent furniture sales transaction report data that has been formatted into the vertical data format.

Table 8. Data Format Vertical

Item	TID_1	TID_2	TID_3	TID_4	TID_5
LPP.32 - Walnut	INVPR2015107801	INVPR2015107844	INVPR2015107876	INVPR2015107883	INVPR2015107896
LPP.22.M - Italian Walnut	INVPR2015107801				
LPP.22 - Beige	INVPR2015107801	INVPR2015107896			

3. Determine the minimum value support. (Minimum Support = 3)
4. Determine the minimum value confidence. (Minimum Confidence = 50%)
5. From the vertical data format can be known candidate itemset 1- (C1). Such data can be seen in the table below.

Table 9. Prospective candidate 1-itemset (C1)

Item	TID_1	TID_2	TID_3	TID_4	TID_5
LPP.32 - Walnut	INVPR2015107801	INVPR2015107844	INVPR2015107876	INVPR2015107883	INVPR2015107896
LPP.22.M - Italian Walnut	INVPR2015107801				
LPP.22 - Beige	INVPR2015107801	INVPR2015107896			
LPC.22.R - Beige	INVPR2015107802	INVPR2015107810	INVPR2015107846	INVPR2015107852	INVPR2015107866
MR03 - Italian Walnut	INVPR2015107802	INVPR2015107811	INVPR2015107820	INVPR2015107837	INVPR2015107839
MB.0606 - Beige	INVPR2015107803	INVPR2015107872			
RTV.0806 - Beige	INVPR2015107803				
BLK.0402.F - Italian Walnut	INVPR2015107804	INVPR2015107848			
BLK.0402 - Beige	INVPR2015107804	INVPR2015107833	INVPR2015107879		

6. Data candidate itemset 1- (C1) that support its value does not meet the minimum value support is removed, it will get frequent 1- itemsets (L1). Such data can be seen in the table below.

Table 10. Frequent 1- itemset (L1)

Item	TID_1	TID_2	TID_3	TID_4	TID_5
LPP.32 - Walnut	INVPR2015107801	INVPR2015107844	INVPR2015107876	INVPR2015107883	INVPR2015107896
LPC.22.R - Beige	INVPR2015107802	INVPR2015107810	INVPR2015107846	INVPR2015107852	INVPR2015107866
MR03 - Italian Walnut	INVPR2015107802	INVPR2015107811	INVPR2015107820	INVPR2015107837	INVPR2015107839
BLK.0402 - Beige	INVPR2015107804	INVPR2015107833	INVPR2015107879		
LHS.0703 - Walnut	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107847	INVPR2015107851
LHS.0701 - Beige	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107861	INVPR2015107870
LPC.32.R - Walnut	INVPR2015107806	INVPR2015107810	INVPR2015107822	INVPR2015107831	INVPR2015107852
RB01 - Walnut	INVPR2015107811	INVPR2015107815	INVPR2015107828	INVPR2015107831	INVPR2015107835
MB.0602 - Beige	INVPR2015107812	INVPR2015107816	INVPR2015107843	INVPR2015107862	INVPR2015107874

7. After getting the data frequent itemset 1-, 2- to obtain candidate itemset (C2) then do a combination between items on L1 based TID_k these items. If the results support the combination produces a value less than the minimum value support, then the itemset is deleted (prune). If the result of the combination generate support value exceeds or equals the minimum value support, then do re-search to get the candidate itemsets 3-, 4- itemset, and so on.

Table 11. The candidate 2-itemset (C2)

Item_1	Item_2	TID_1	TID_2	TID_3	TID_4
LPP.32 - Walnut	MR03 - Italian Walnut	INVPR2015107896			
LPP.32 - Walnut	MR02 - Italian Walnut	INVPR2015107883			
LPP.32 - Walnut	LPP.22 - Italian Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883	
LPC.22.R - Beige	MR03 - Italian Walnut	INVPR2015107802			
LPC.22.R - Beige	LHS.0703 - Walnut	INVPR2015107846			
LPC.22.R - Beige	LHS.0701 - Beige	INVPR2015107846			
LPC.22.R - Beige	LPC.32.R - Walnut	INVPR2015107810	INVPR2015107852	INVPR2015107884	
LPC.22.R - Beige	RB01 - Walnut	INVPR2015107866			
LPC.22.R - Beige	MT0901 - Walnut	INVPR2015107890			
LPC.22.R - Beige	RTV.0802 - Italian Walnut	INVPR2015107866			
LPC.22.R - Beige	LHS.0703 - Italian Walnut	INVPR2015107878			

Table 12. Frequent 2-itemset (L2)

Item_1	Item_2	TID_1	TID_2	TID_3	TID_4	Support
LPP.32 - Walnut	LPP.22 - Italian Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3
LPC.22.R - Beige	LPC.32.R - Walnut	INVPR2015107810	INVPR2015107852	INVPR2015107884		3
MR03 - Italian Walnut	MR02 - Italian Walnut	INVPR2015107820	INVPR2015107839	INVPR2015107845	INVPR2015107864	4
LHS.0703 - Walnut	LHS.0701 - Beige	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4
MB.0602 - Beige	MB.0603 - Beige	INVPR2015107843	INVPR2015107862	INVPR2015107893		3

8. Perform the 7th in berulangan on each item to not generate new itemset. The results can be seen in the table below.

Table 13. Frequent 2-itemsets (L2)

Item_1	Item_2	TID_1	TID_2	TID_3	TID_4	Support
LPP.32 - Walnut	LPP.22 - Italian Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3
LPC.22.R - Beige	LPC.32.R - Walnut	INVPR2015107810	INVPR2015107852	INVPR2015107884		3
MR03 - Italian Walnut	MR02 - Italian Walnut	INVPR2015107820	INVPR2015107839	INVPR2015107845	INVPR2015107864	4
LHS.0703 - Walnut	LHS.0701 - Beige	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4
MB.0602 - Beige	MB.0603 - Beige	INVPR2015107843	INVPR2015107862	INVPR2015107893		3

9. From the result of the formation of itemset meets the minimum support value, then the next step is to calculate the confidence value for each itemset generated.

Table 14. Results Value Calculation Confidence

Itemset			TID_1	TID_2	TID_3	TID_4	Support	Confidence (%)
LPP.32 - Walnut	→	LPP.22 - Italian Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3	60
LPC.22.R - Beige	→	LPC.32.R - Walnut	INVPR2015107810	INVPR2015107852	INVPR2015107884		3	37,5
MR03 - Italian Walnut	→	MR02 - Italian Walnut	INVPR2015107820	INVPR2015107839	INVPR2015107845	INVPR2015107864	4	36,36
LHS.0703 - Walnut	→	LHS.0701 - Beige	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4	66,67
MB.0602 - Beige	→	MB.0603 - Beige	INVPR2015107843	INVPR2015107862	INVPR2015107893		3	50
LPP.22 - Italian Walnut	→	LPP.32 - Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3	100
LPC.32.R - Walnut	→	LPC.22.R - Beige	INVPR2015107810	INVPR2015107852	INVPR2015107884		3	42,85
MR02 - Italian Walnut	→	MR03 - Italian Walnut	INVPR2015107820	INVPR2015107839	INVPR2015107845	INVPR2015107864	4	44,44
LHS.0701 - Beige	→	LHS.0703 - Walnut	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4	66,67
MB.0603 - Beige	→	MB.0602 - Beige	INVPR2015107843	INVPR2015107862	INVPR2015107893		3	60

10. After getting the confidence value calculation result data of each itemset, the next step is to delete the data that does not meet the minimum value confidence.

Table 15. Data that Meets Association Rules

Itemset			TID_1	TID_2	TID_3	TID_4	Support	Confidence (%)
LPP.32 - Walnut	→	LPP.22 - Italian Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3	60
LHS.0703 - Walnut	→	LHS.0701 - Beige	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4	66,67
MB.0602 - Beige	→	MB.0603 - Beige	INVPR2015107843	INVPR2015107862	INVPR2015107893		3	50
LPP.22 - Italian Walnut	→	LPP.32 - Walnut	INVPR2015107844	INVPR2015107876	INVPR2015107883		3	100
LHS.0701 - Beige	→	LHS.0703 - Walnut	INVPR2015107806	INVPR2015107842	INVPR2015107846	INVPR2015107898	4	66,67

2.4. Context Diagram

Context diagram illustrates the process input and output data on the system. Here is a context diagram of applications built :

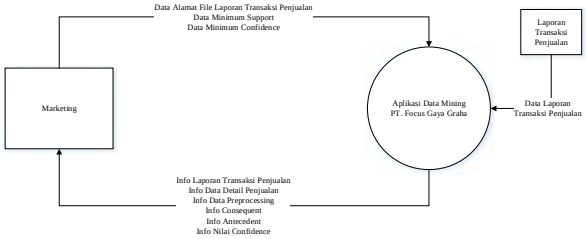


Figure 3. Context Diagram

2.5. Data Flow Diagram(DFD)

DFD is a logical model of data or processes are created to illustrate where the origin of the data and where the purpose of the data coming out of the system. Here is the DFD level 1 of a system built :

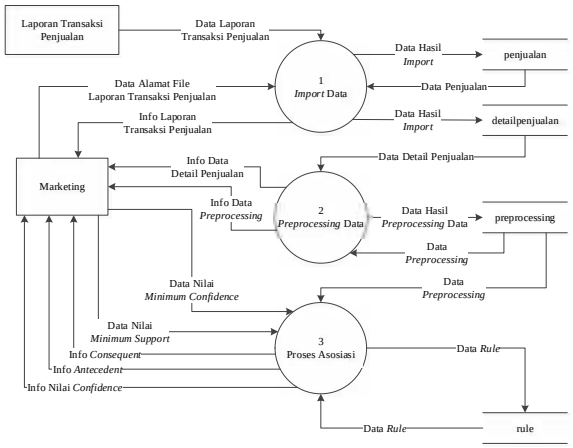


Figure 4. DFD level 1

Here is the DFD level 2 of the Import Data :

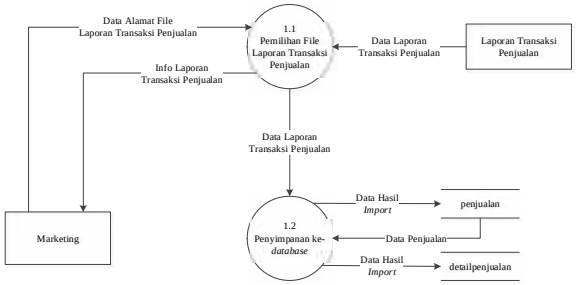


Figure 5. DFD level 2 for the Import Data

Here is the DFD level 2 process data preprocessing :

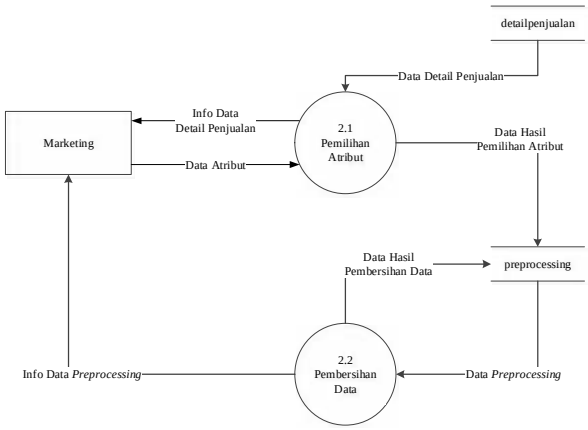


Figure 6. DFD level 2 for process Preprocessing Data

Here is the DFD level 2 for the Association :

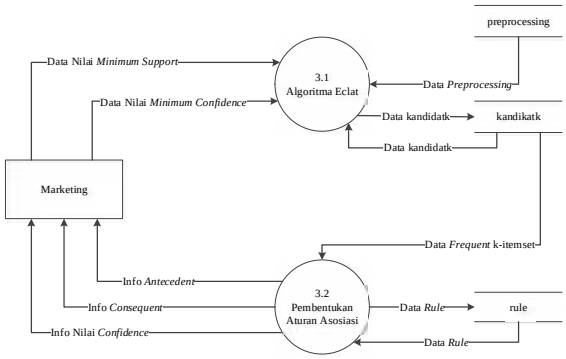


Figure 7. DFD level 2 for the process of Association

Here is the menu structure of the applications built :

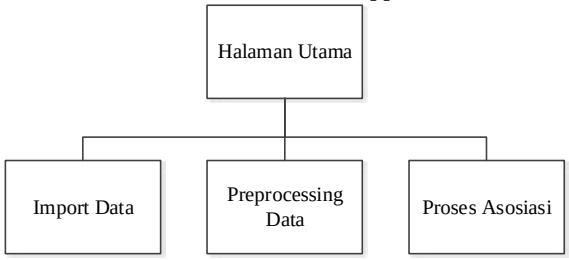


Figure 8. The menu structure

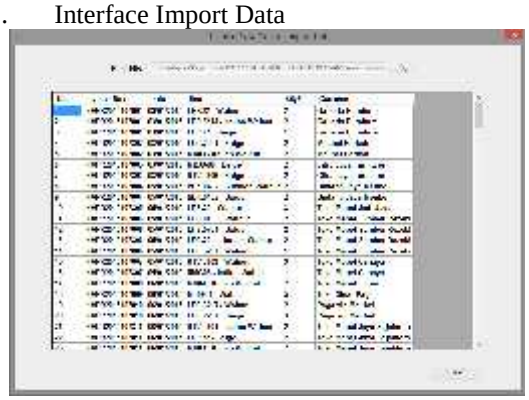
2.6 Interface Implementation

a. Interface Home



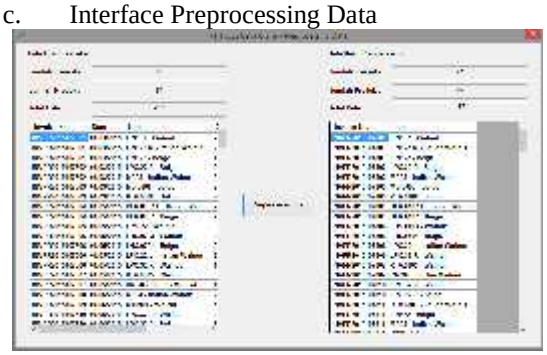
Gambar 9. Interface Home

b. Interface Import Data



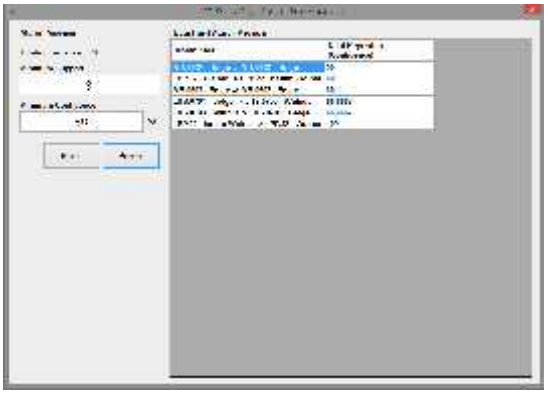
Gambar 10. Interface Import Data

c. Interface Preprocessing Data



Gambar 11. Interface Preprocessing Data

d. Interface Proses Asosiasi



Gambar12. Interface Proses Asosiasi

3. RESEARCH RESULT

The research results are tested by testing the software that has been built is generating the desired data and in accordance with the results of applying the method of association rule algorithms Eclat.

This test uses the sales transaction data sample report period January-February 2015 by value minimum support 3 and the minimum value 50% confidence. The resulting association rules system built are :

Itemset			Confidence (%)
MB.0602 - Beige	→	MB.0603 - Beige	50
LPP.32 - Walnut	→	LPP.22 - Italian Walnut	60
MB.0603 - Beige	→	MB.0602 - Beige	60
LHS.0701 - Beige	→	LHS.0703 - Walnut	66,67
LHS.0703 - Walnut	→	LHS.0701 - Beige	66,67
LPP.22 - Italian Walnut	→	LPP.32 - Walnut	100

Figure 13. The resulting association rules system

Based on the test results of samples above it can be concluded that the minimum value support and minimum confidence greatly affects the number of rules that are formed. Therefore for the minimum value better support small value, while the minimum value better confidence of great value because if such a rule would produce a more varied with high certainty value.

4. CLOSING

4.1. Conclusion

Based on the results of the implementation and testing has been done on the applications that have been built, it can be concluded that an application built to help PT. Focus Style Graha to determine consumer purchasing patterns, so it can help companies to determine which products will be produced.

4.2. Suggestion

The suggestions for further application development is as follows:

1. Results associative rule (rule) that is produced can be grouped by sales period and sorted based on the value of his confidence in order to facilitate the selection of the products which will go into the production process.
2. Optimization in terms of time and memory, as more and more data is processed, the more also time and memory required to form itemset..

BIBLIOGRAPHY

[1] I. Pramudiono, Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data, 2007.

[2] B. Santosa, Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis, Yogyakarta: Graha Ilmu, 2007.

[3] A. Dodiet, Penelitian Deskriptif, 2009.

[4] D. T. Larose, DISCOVERING KNOWLEDGE IN DATA : An Introduction to DATA MINING, Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.

[5] I. Sommerville, SOFTWARE ENGINEERING 9 Edition, PEARSON, 2011.

[6] Fathayansyah, Buku Teks Komputer : Basis Data, 5th ed., Bandung: Informatika, 2004.

[7] Waljiyanto, Sistem Basis Data: Analisis dan Pemodelan Data, 1st ed., Yogyakarta: Graha Ilmu, 2003.

[8] J. Simarmata dan I. Prayudi, Basis Data, Yogyakarta: Andi Offset, 2006.

[9] Kusrini dan E. T. Luthfi, ALGORITMA DATA MINING Ed : 1, Yogyakarta: Andi, 2009.

[10] D. T. Larose, DATA MINING METHODS AND MODELS, New Jersey: John Wiley & Sons, Inc, 2006.

[11] D. Ulmer, “Mining an Online Auctions Data Warehouse,” The Mid-Atlantic Student Workshop on Programming Languages and Systems, vol. 19, 2002.

[12] J. Han, M. Kamber dan J. Pei, DATA MINING : Concepts and Techniques 3rd Edition, USA: Morgan Kaufmann Publishers, 2012.

[13] A. B. b. Ladjamudin, Rekayasa Perangkat Lunak, Yogyakarta: Graha Ilmu, 2006.

[14] A. Kadir, Dasar Aplikasi Database MySQL - Delphi, 2nd ed., Yogyakarta: Andi, 2004.