DESIGN AND MAKE OF DATA MINING MARKET BASKET ANALYSIS APLICATION AT DE JOGLO RESTAURANT

Nadia Andreani Dwiyono
26406019

Background

- De Joglo Restaurant have been using computerized system in any sales transaction, so the number of transactions can be analyzed to provide useful knowledge for knowing customer's order habits. So, the restaurant needs an application which can provide useful information for users

Problem Definition

- How to find patterns association between the different menus which ordered by customer.

Purpose

- Design and make an application to help manager to gain important informations about the food menus are often ordered by customers.

Scope

Input:

- The data that being analyzed for testing are sales data transaction in one year.
- Processing data is from sales details and menu table in a particular period.

Scope (continue)

Process:

- This application used Market Basket Analysis method.
- The results of input will be grouped by order number, then determined the minimum support and then processed using the FP-Tree algorithm.

Scope (continue)

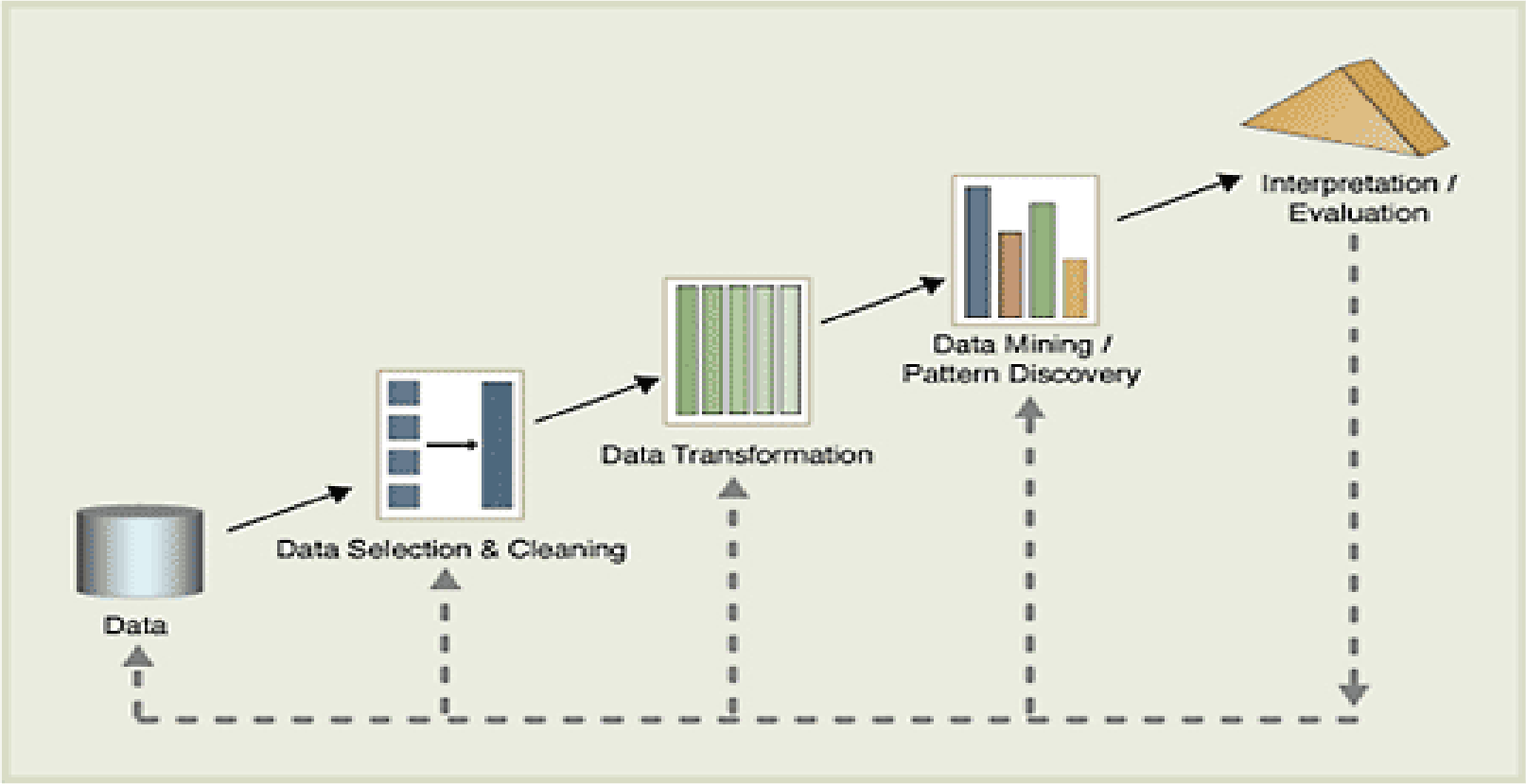
Output:

- The results from the processed tables will be presented in association rules and will be displayed for user information in the table and graph form.
- The program will only provide information to help manager in the restaurant, not offering problem solution.

Data Mining

- Data mining is a process to discover interesting knowledge from large amounts of data which stored in the database (Han, Kamber, 2001).
- Data Mining is also often said by many people as a synonym of Knowledge Discovery in Databases or KDD.

Stages in Data Mining (KDD)



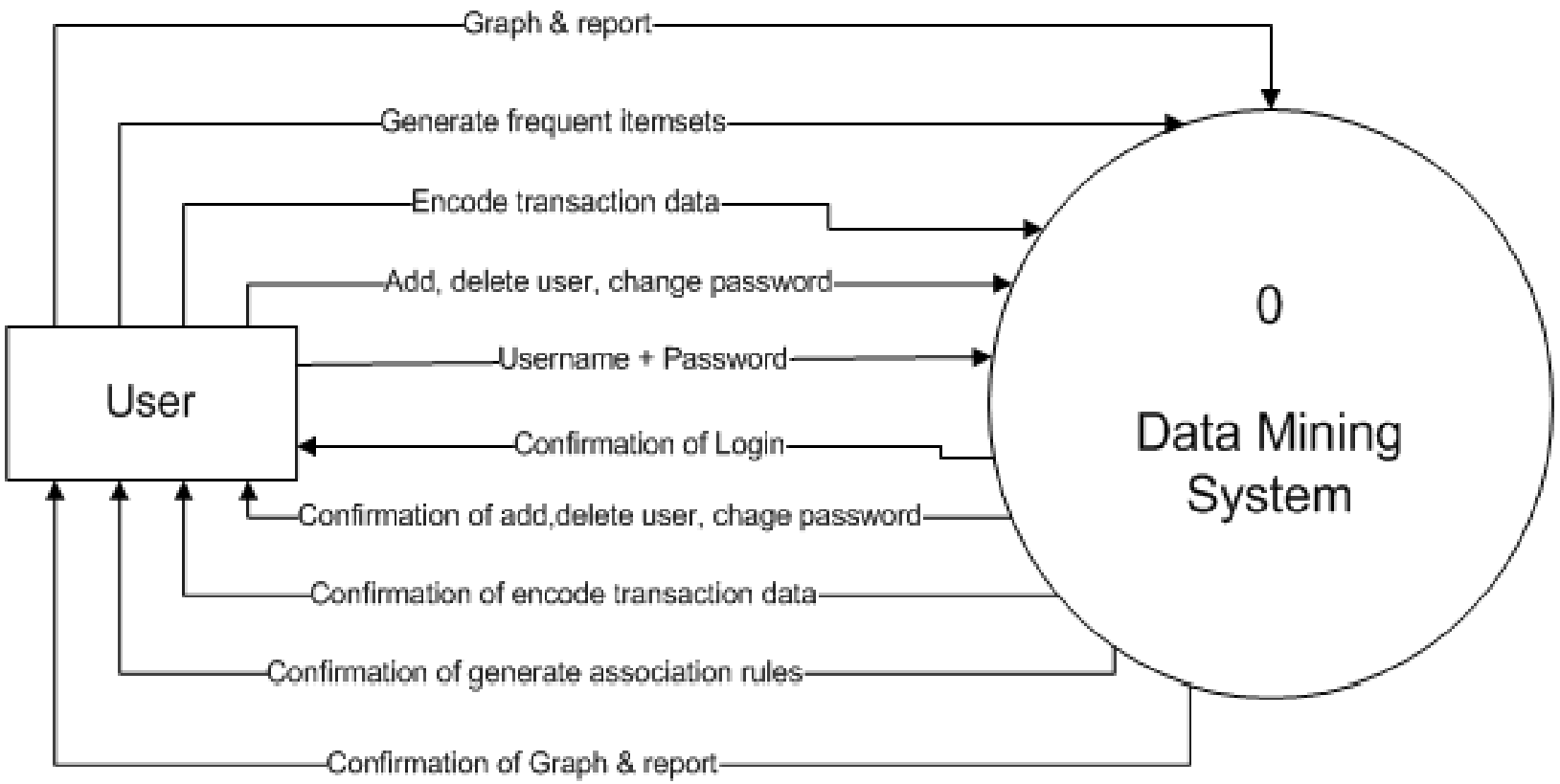
Market Basket Analysis

- *Market Basket Analysis* is interested in identifying which products tend to be purchased together (Olson, Yong, 2007).
- This technique is also called Association rule analysis which is another way to do data mining.

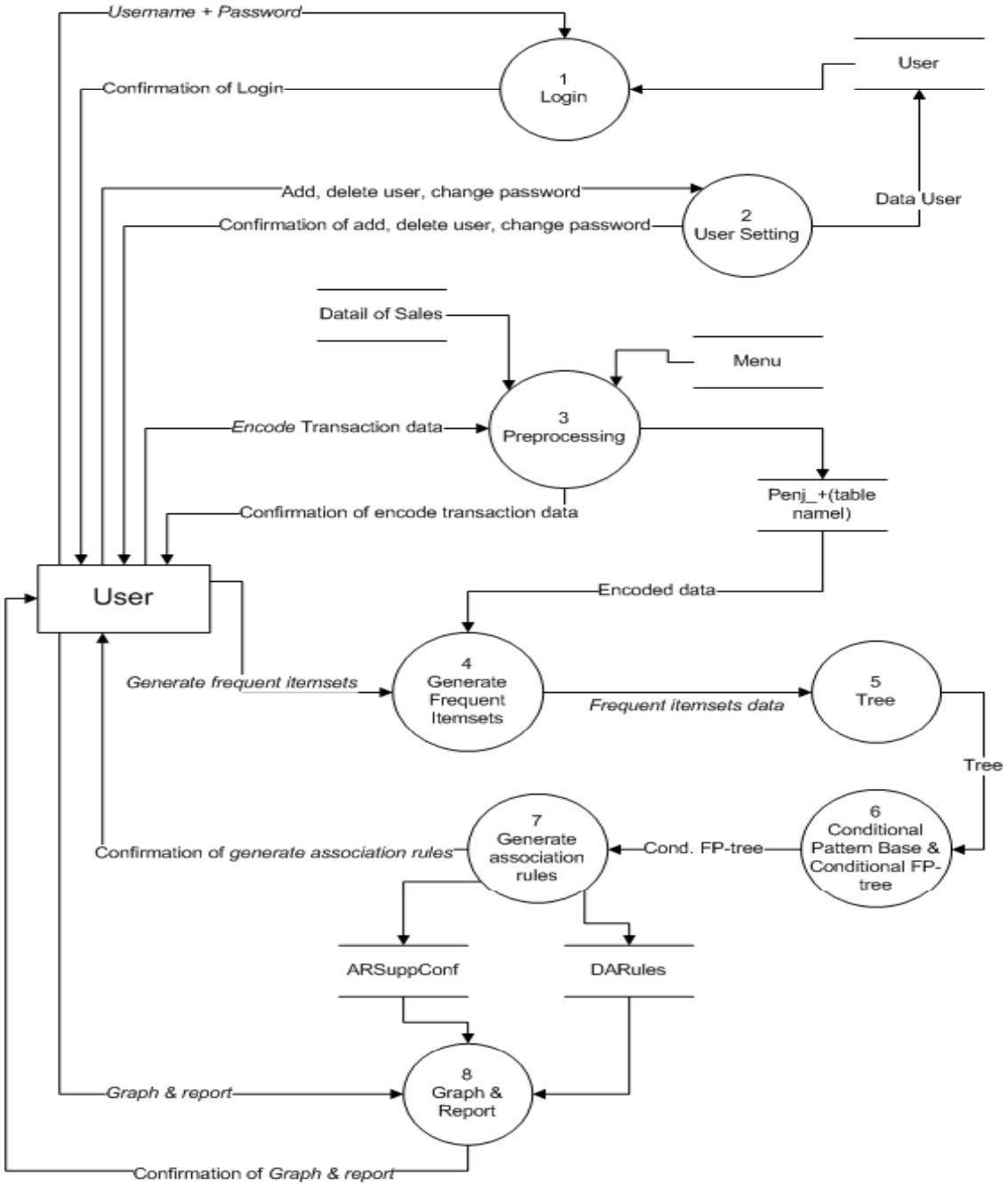
FP-TREE

- FP-Tree is an algorithm used to find association patterns in transaction databases.
- Each node in the FP-Tree contains three important information. The informations are:
  - Label item
  - Count
  - Link

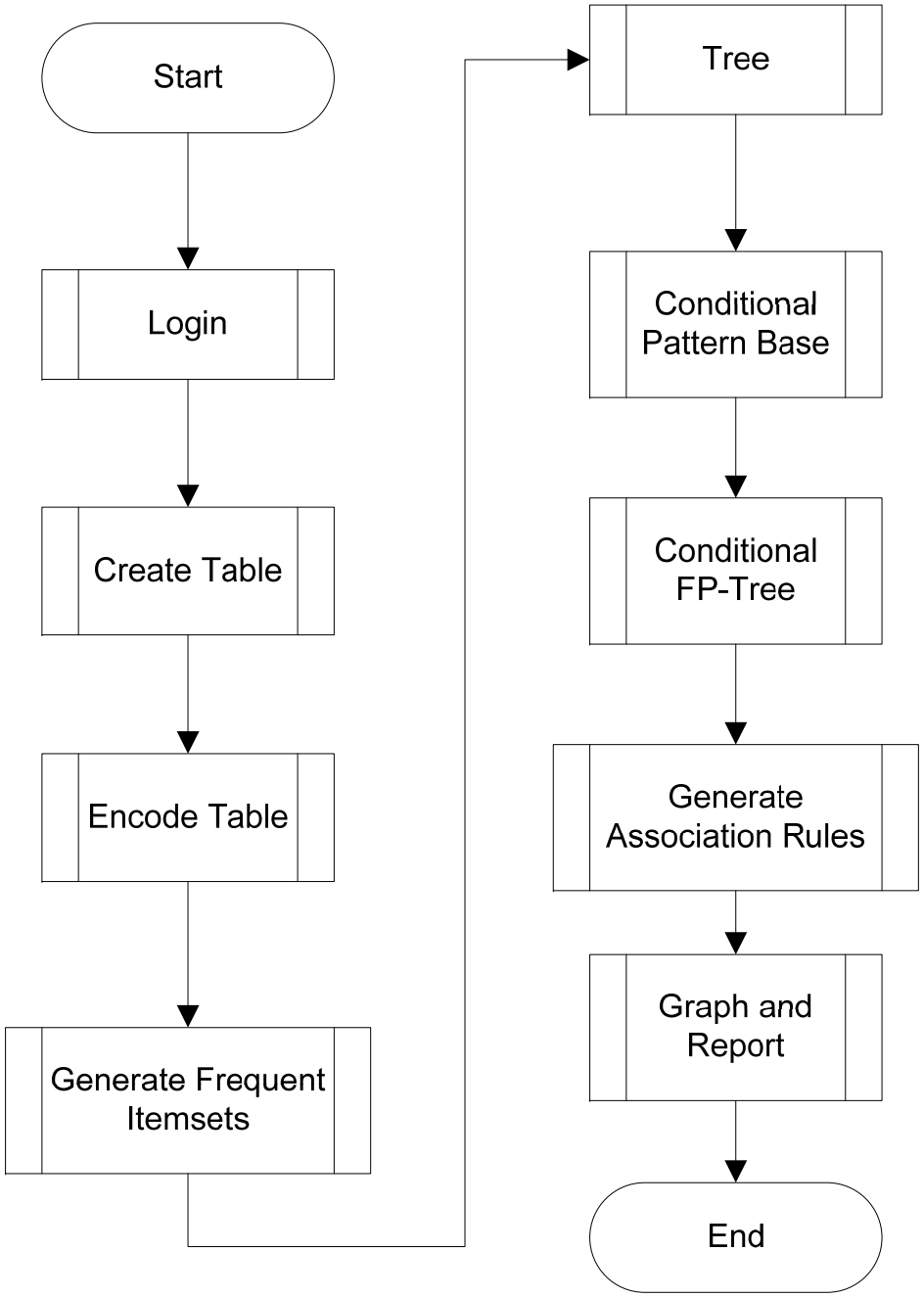
Context Diagram

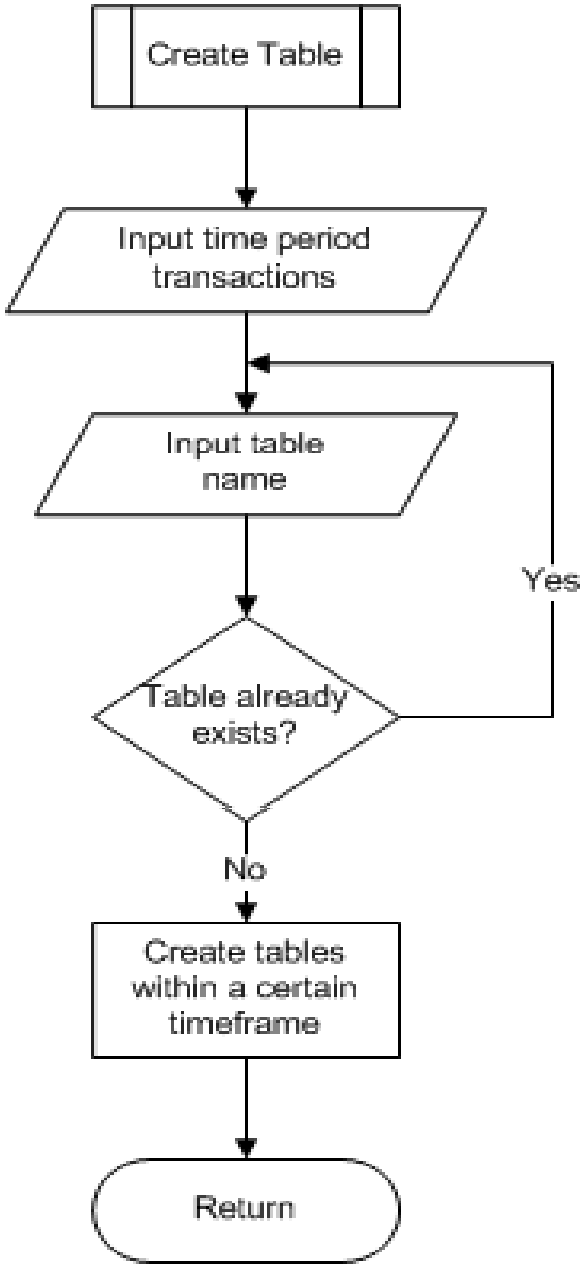


DFD Level 0


Username + Password
Confirmation of Login
1
Login
User
Add, delete user, change password
Confirmation of add, delete user, change password
2
User Setting
Data User
Datail of Sales
Menu
Encode Transaction data
3
Preprocessing
Confirmation of encode transaction data
Penj_+(table
name)
User
Encoded data
Generate frequent itemsets
4
Generate
Frequent
Itemsets
Frequent itemsets data
5
Tree
Tree
Confirmation of generate association rules
7
Generate
association
rules
Cond. FP-tree
6
Conditional
Pattern Base &
Conditional FP-
tree
ARSuppConf
DARules
Graph & report
8
Graph &
Report
Confirmation of Graph & report

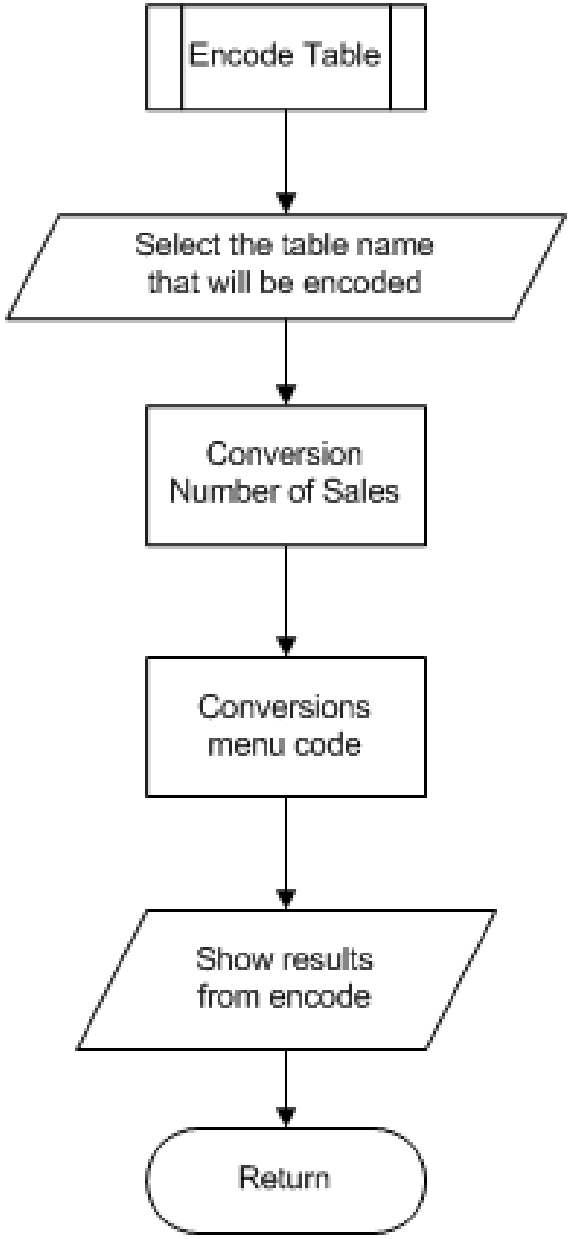
Flowchart


Start
Login
Create Table
Encode Table
Generate Frequent Itemsets
Tree
Conditional Pattern Base
Conditional FP-Tree
Generate Association Rules
Graph and Report
End

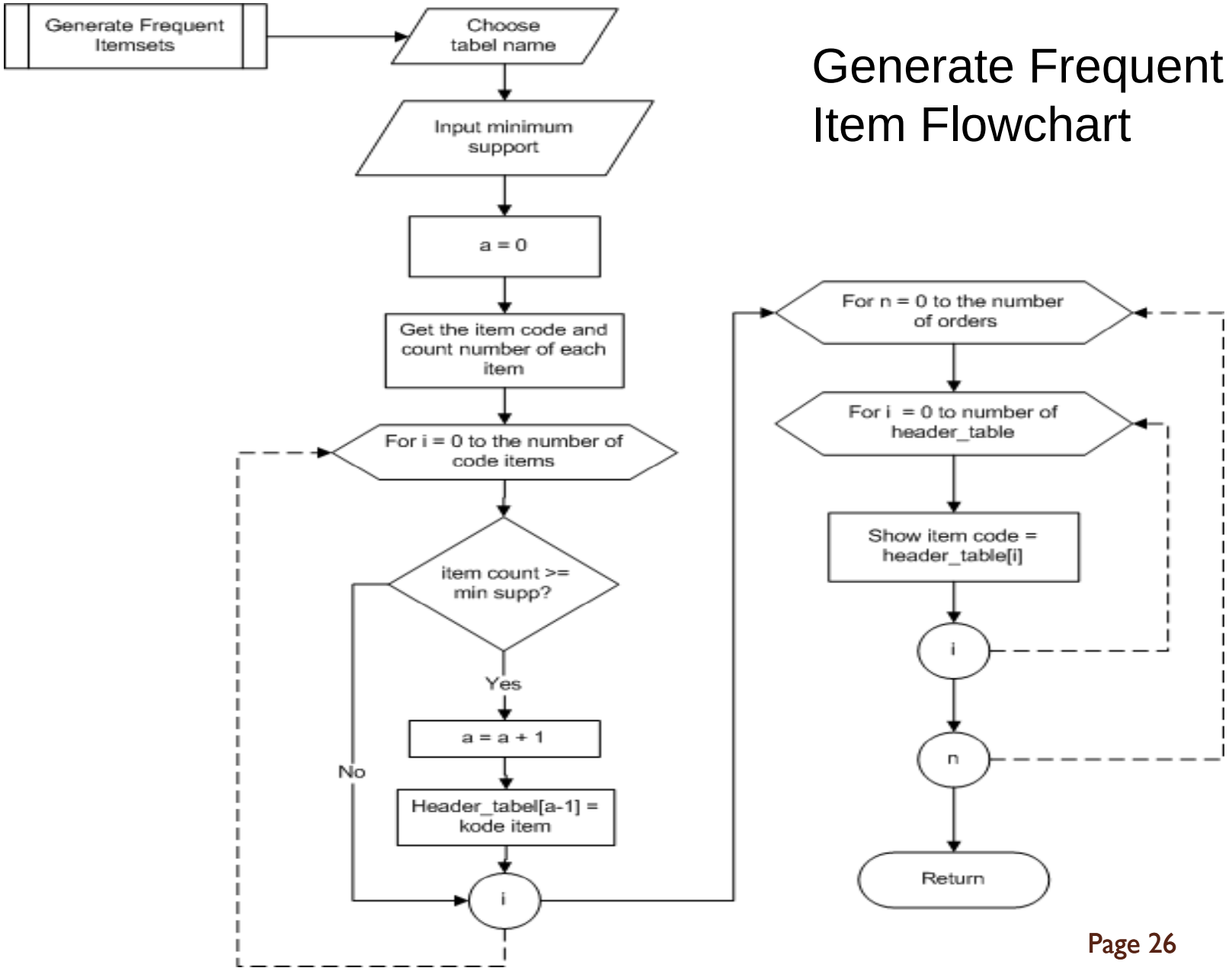


Create Table Flowchart

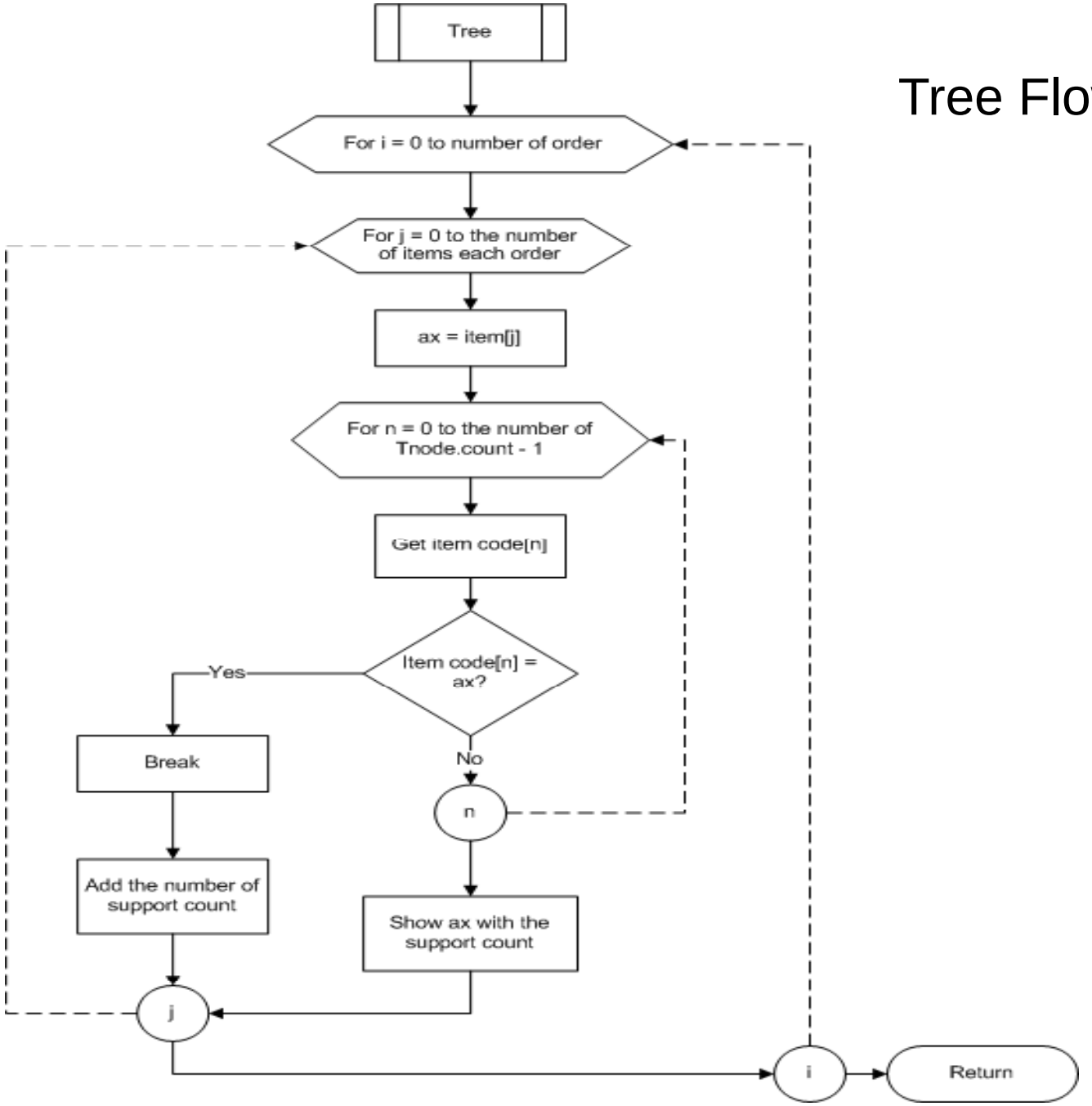
Encode Table Flowchart



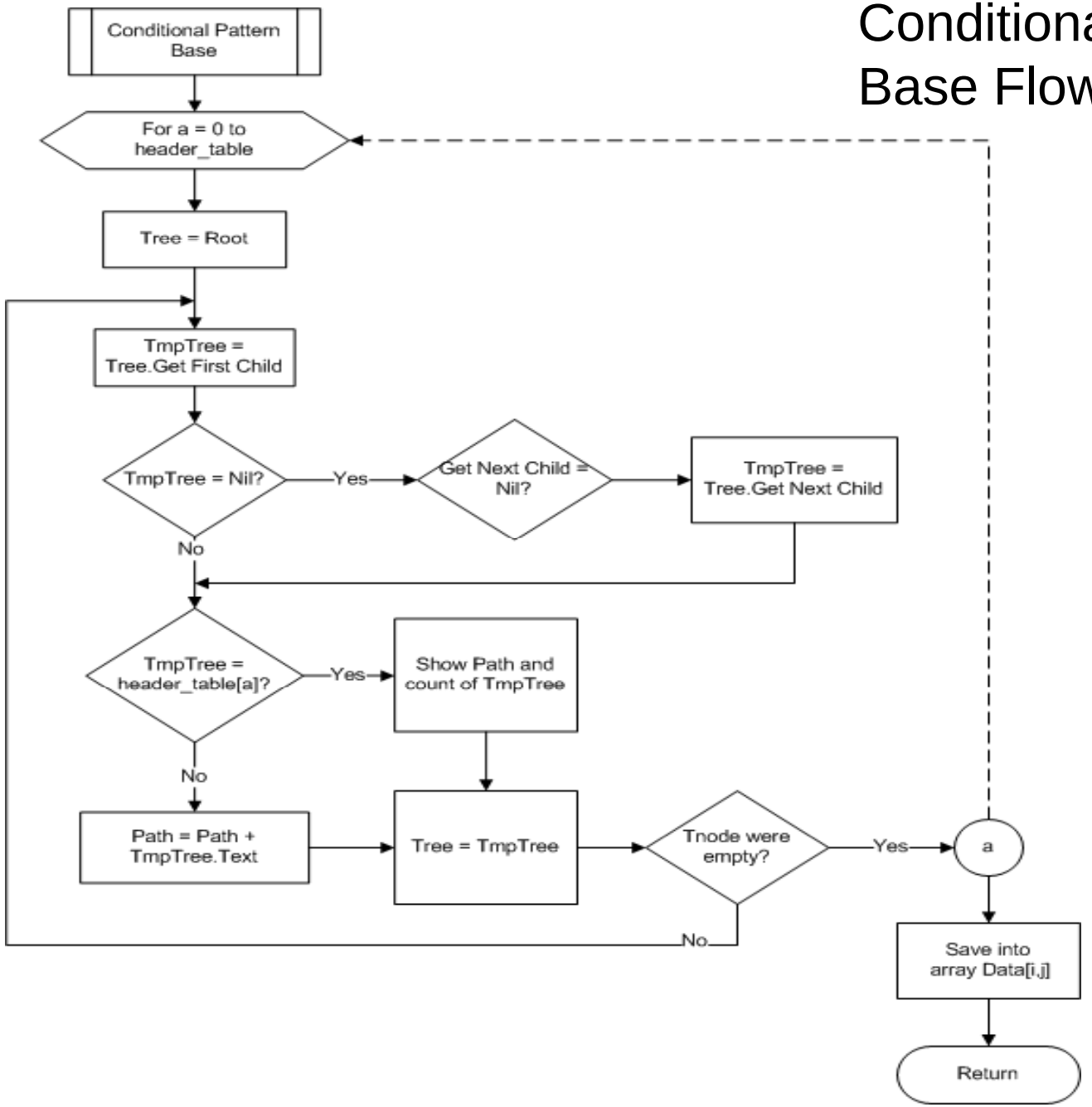
Generate Frequent Item Flowchart



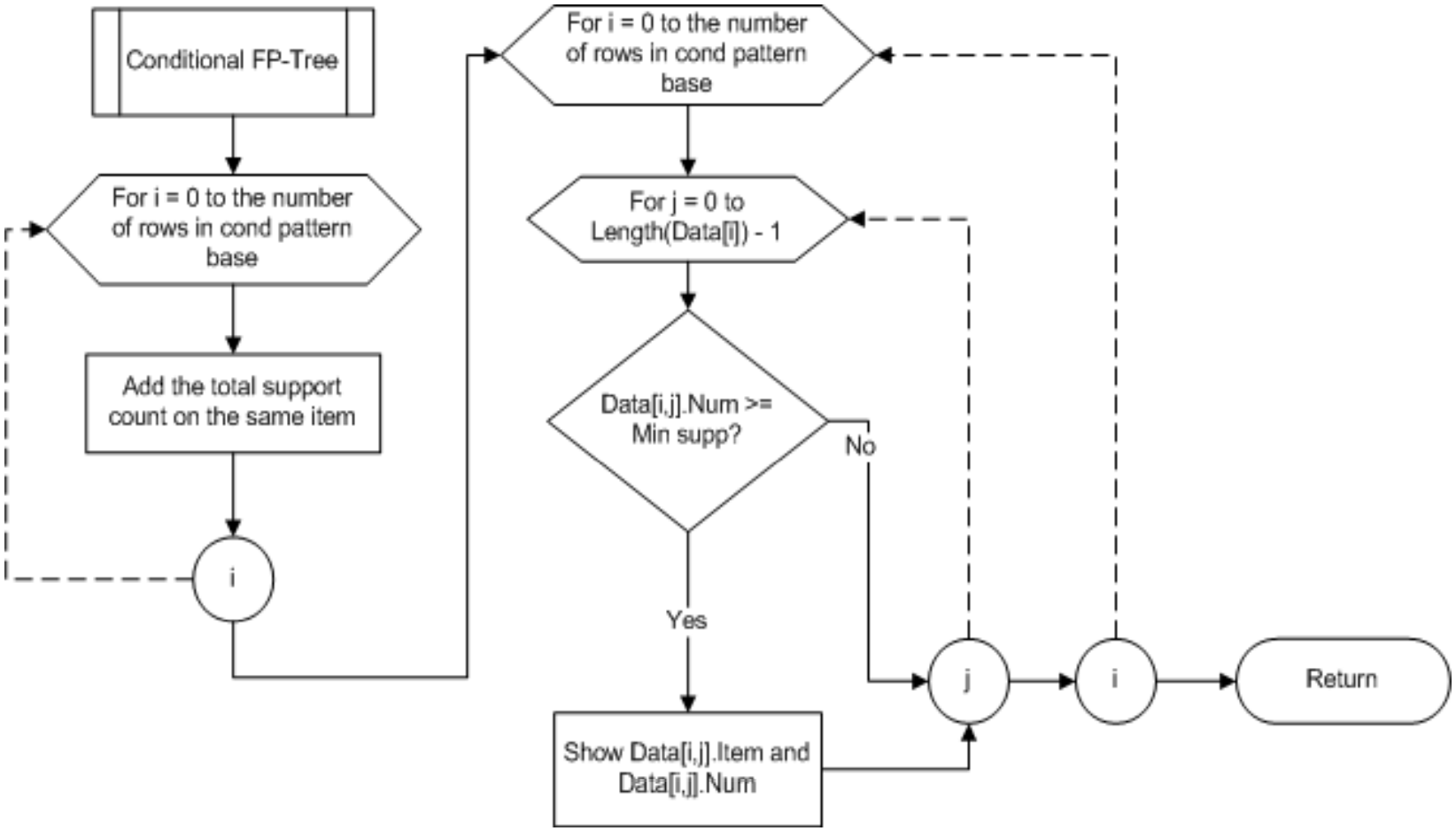
Tree Flowchart



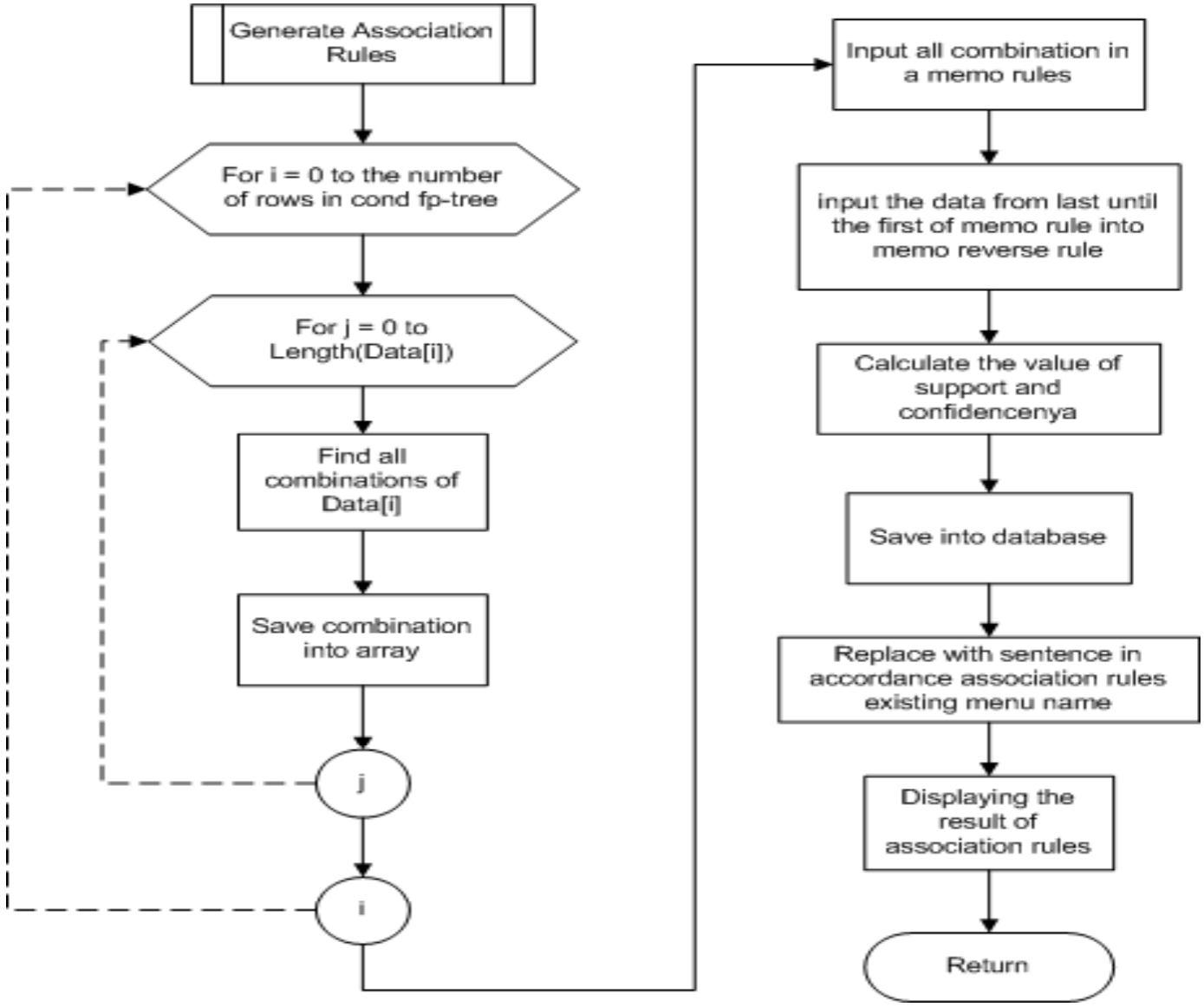
Conditional Pattern Base Flowchart



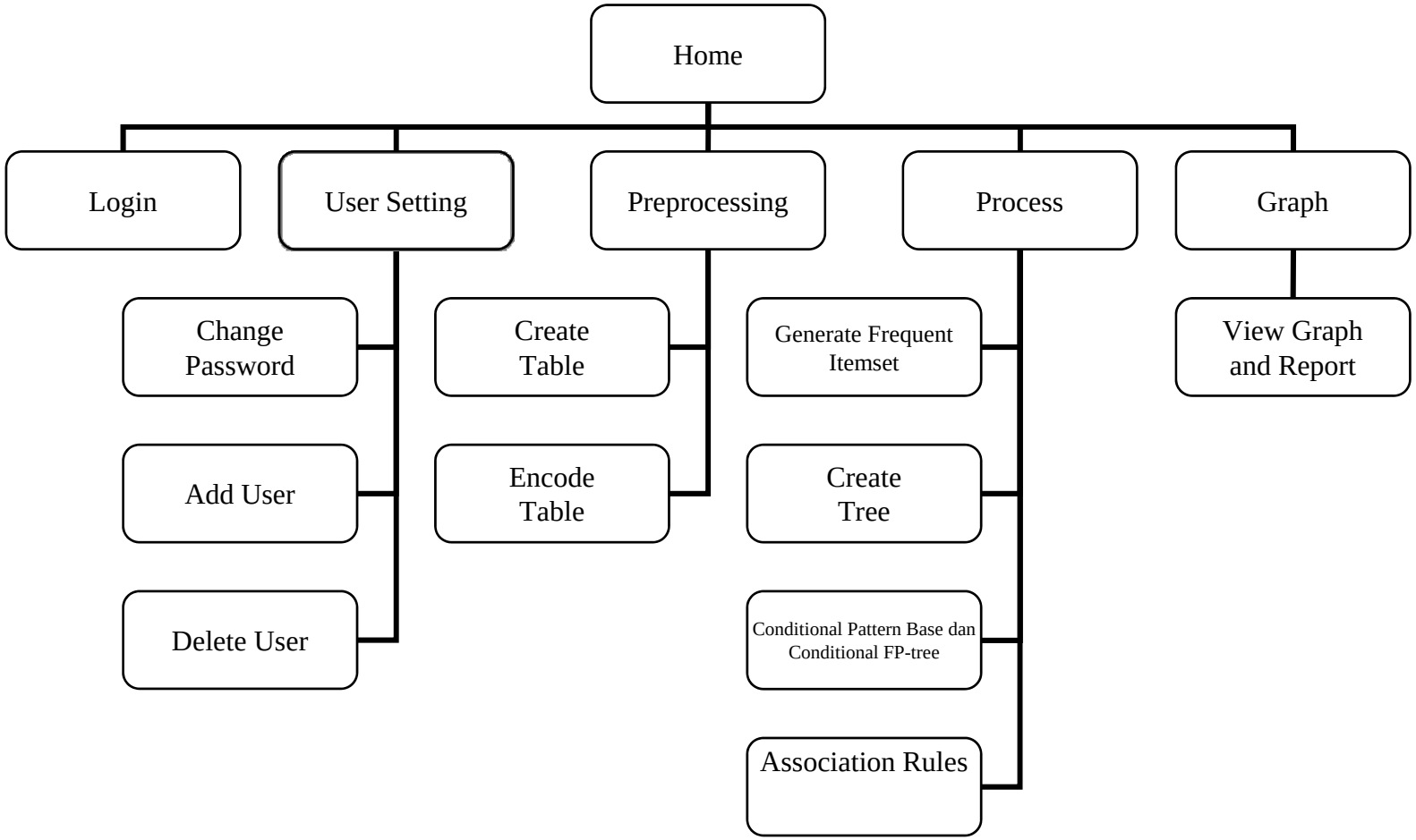
Conditional FP-Tree Flowchart



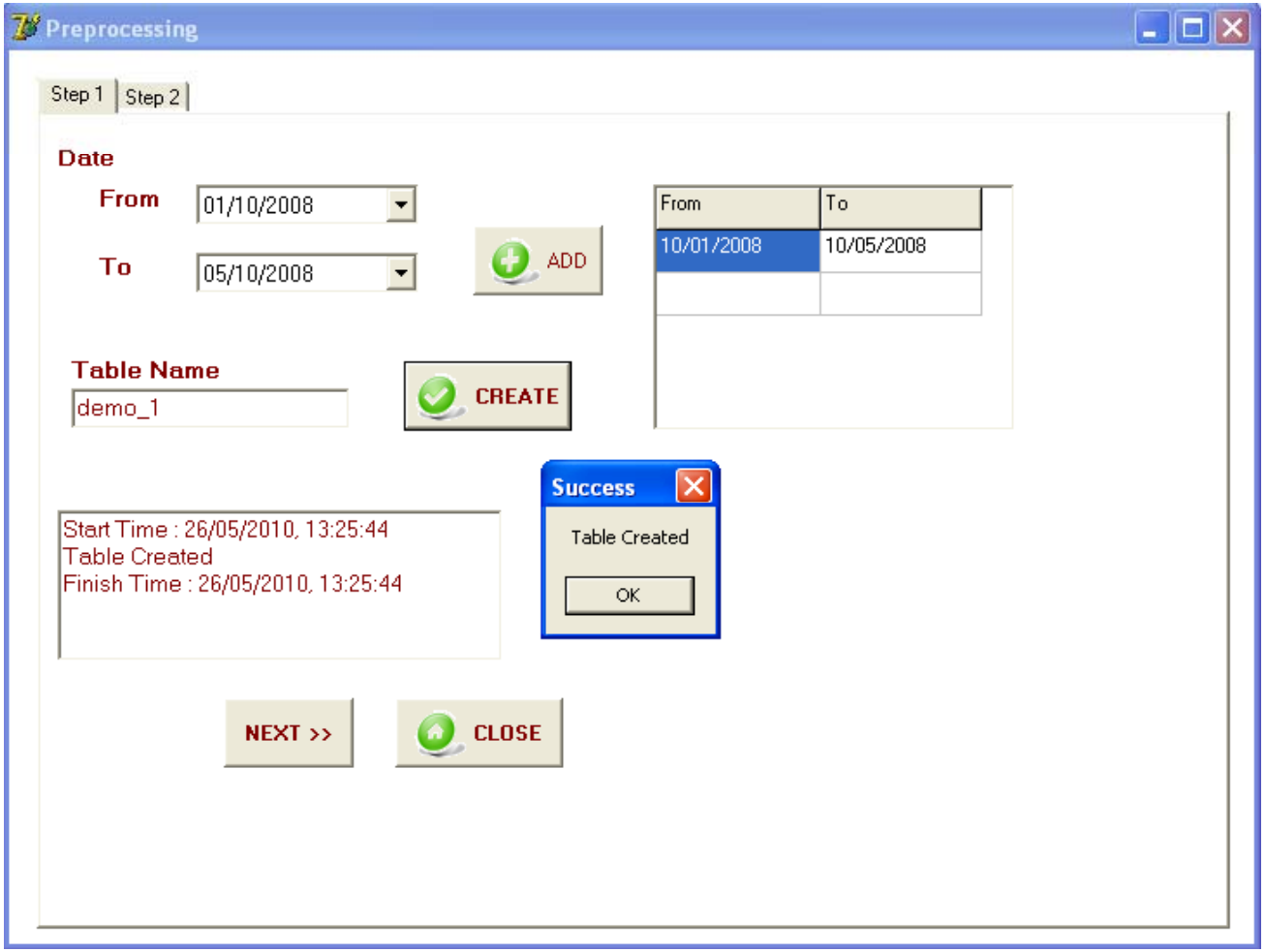
Generate Association Rules Flowchart



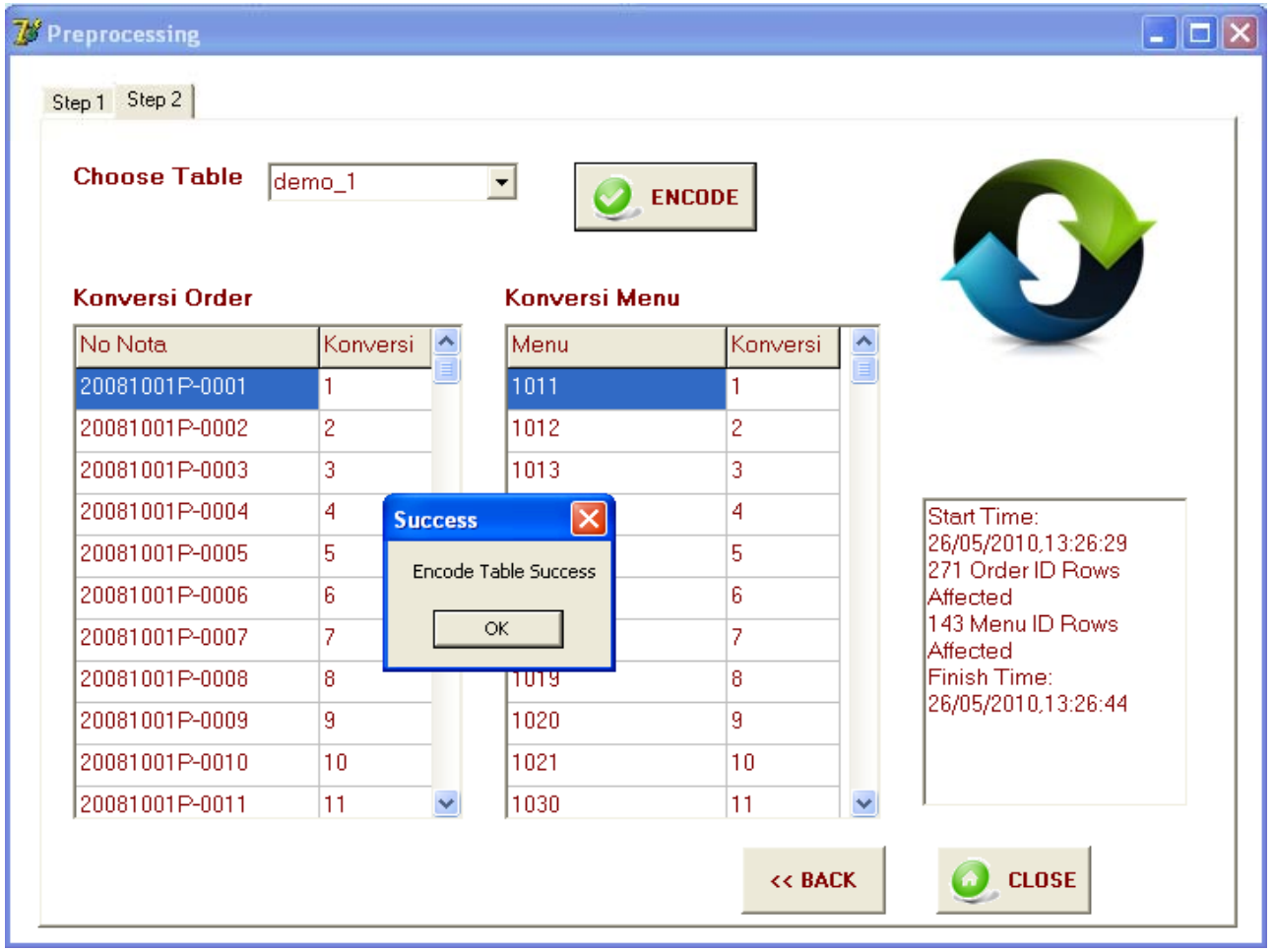
System Design (Menu)



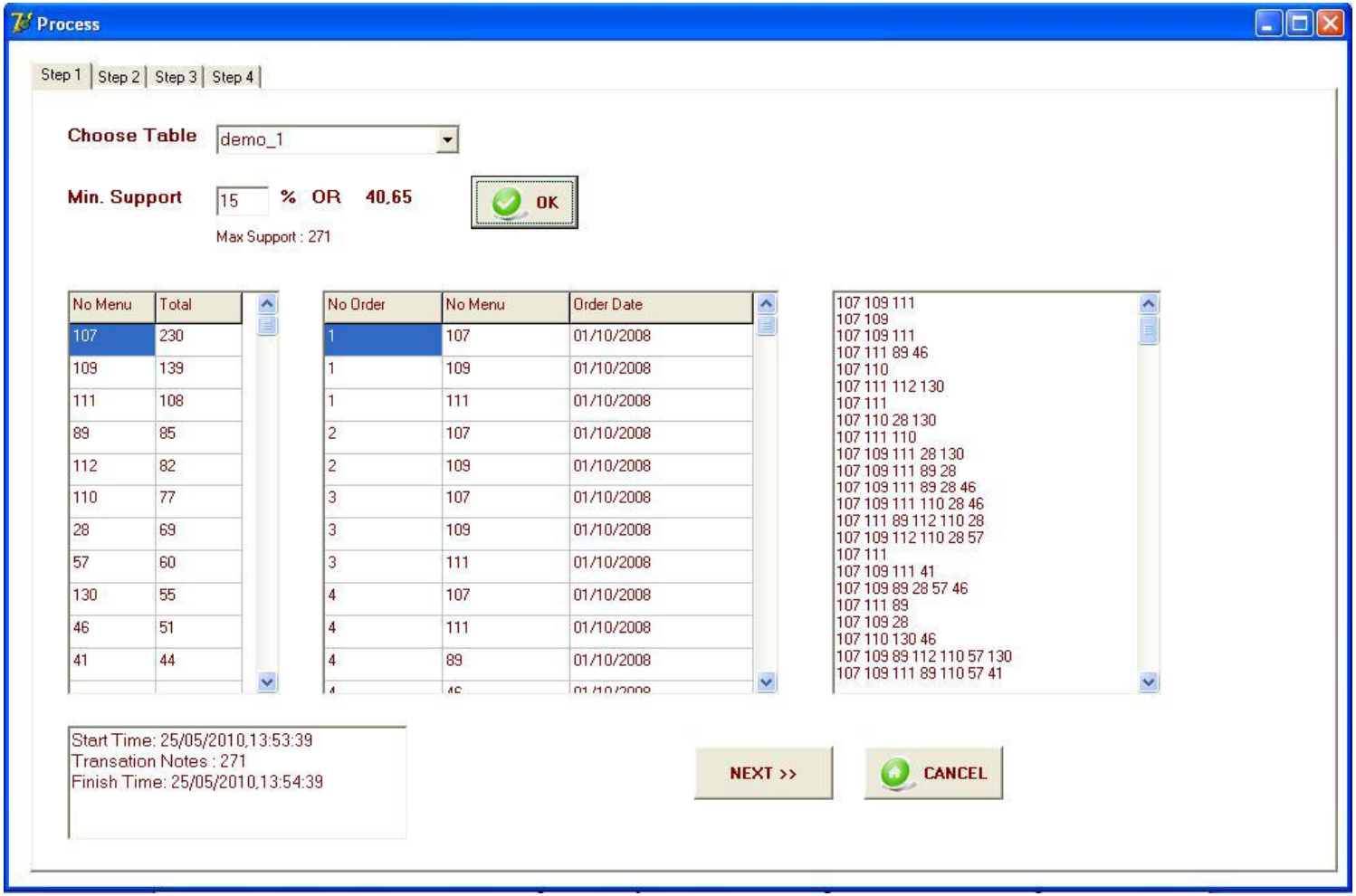
Preprocessing (Create Table)



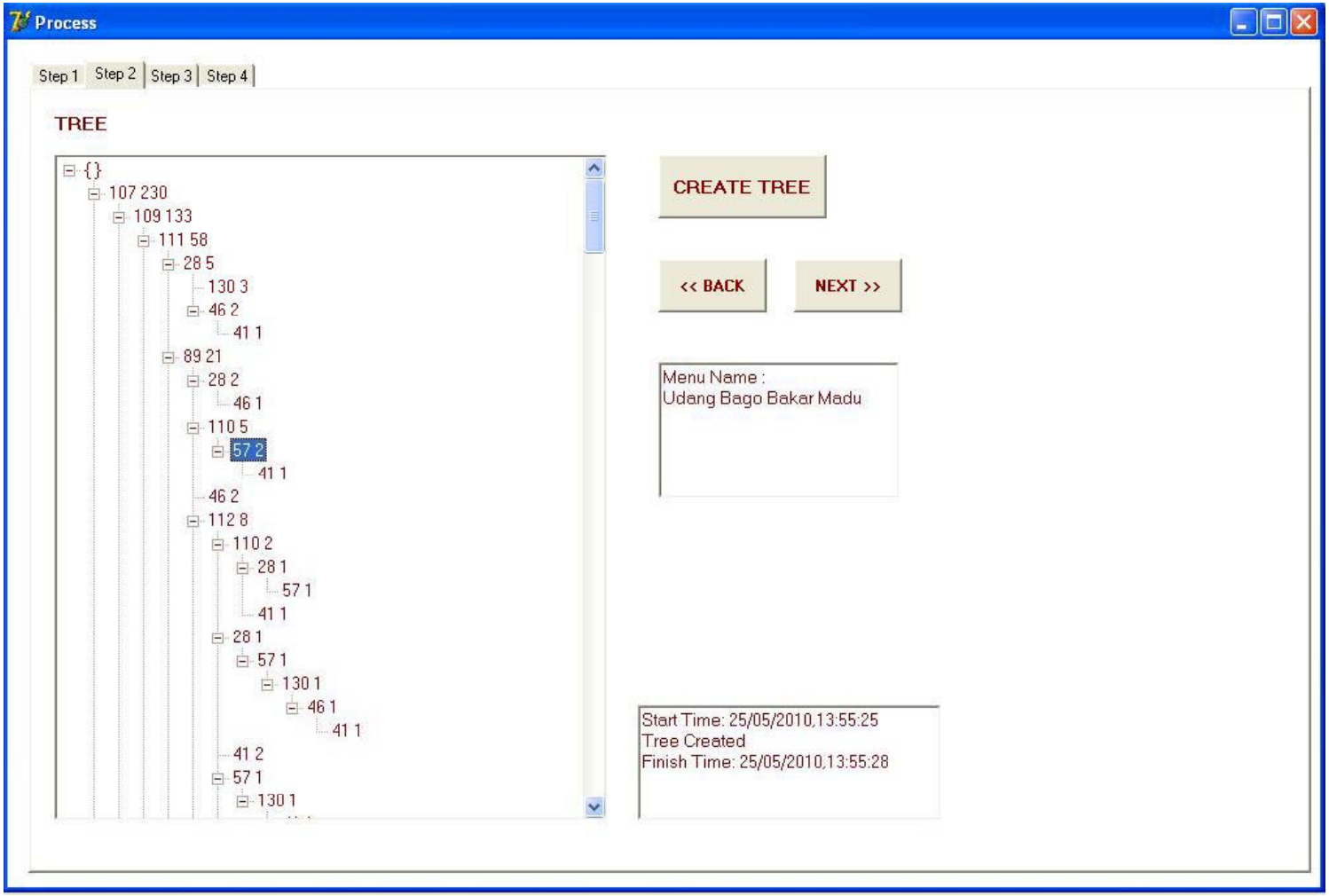
Preprocessing (Encode Table)



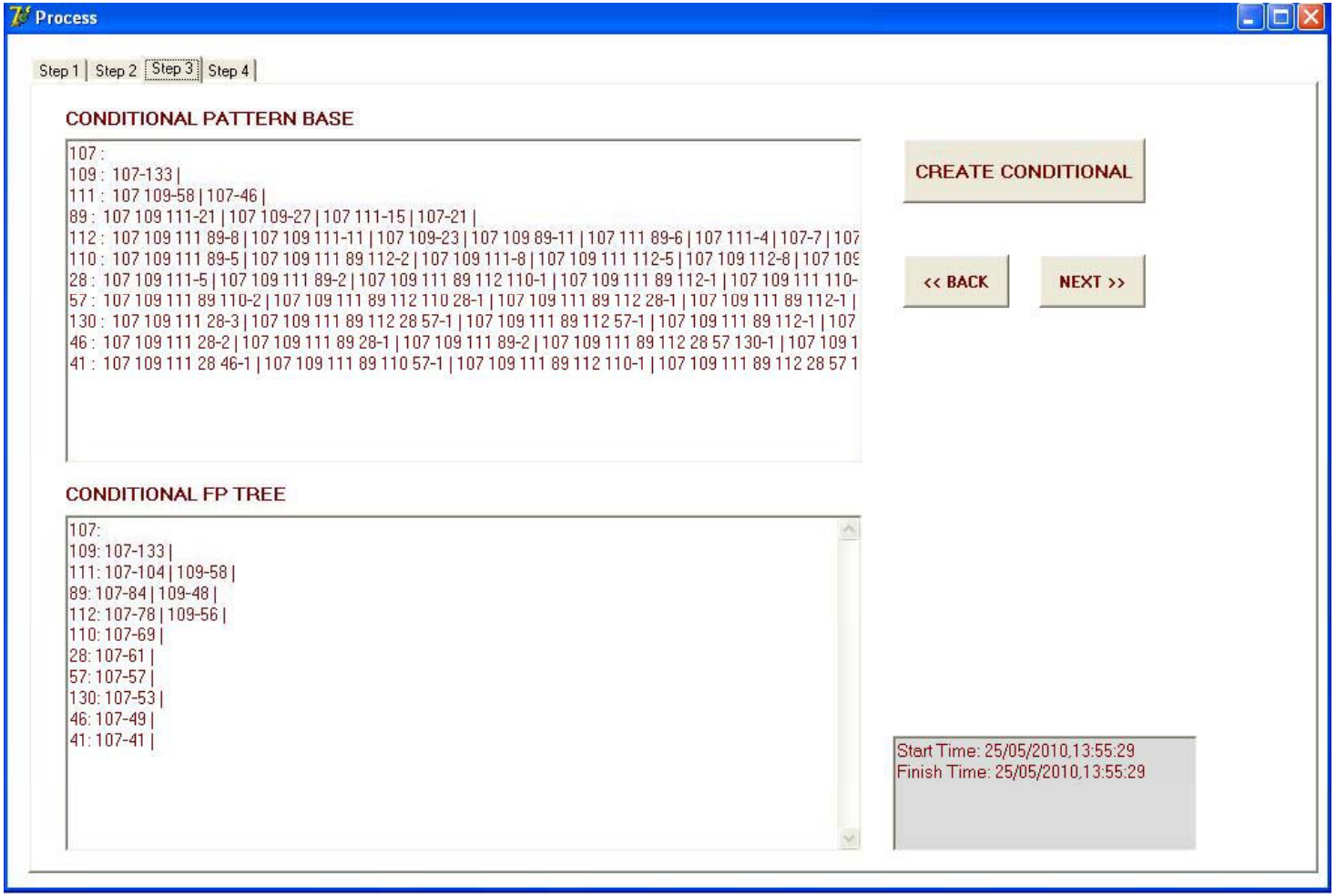
Generate Frequent Item



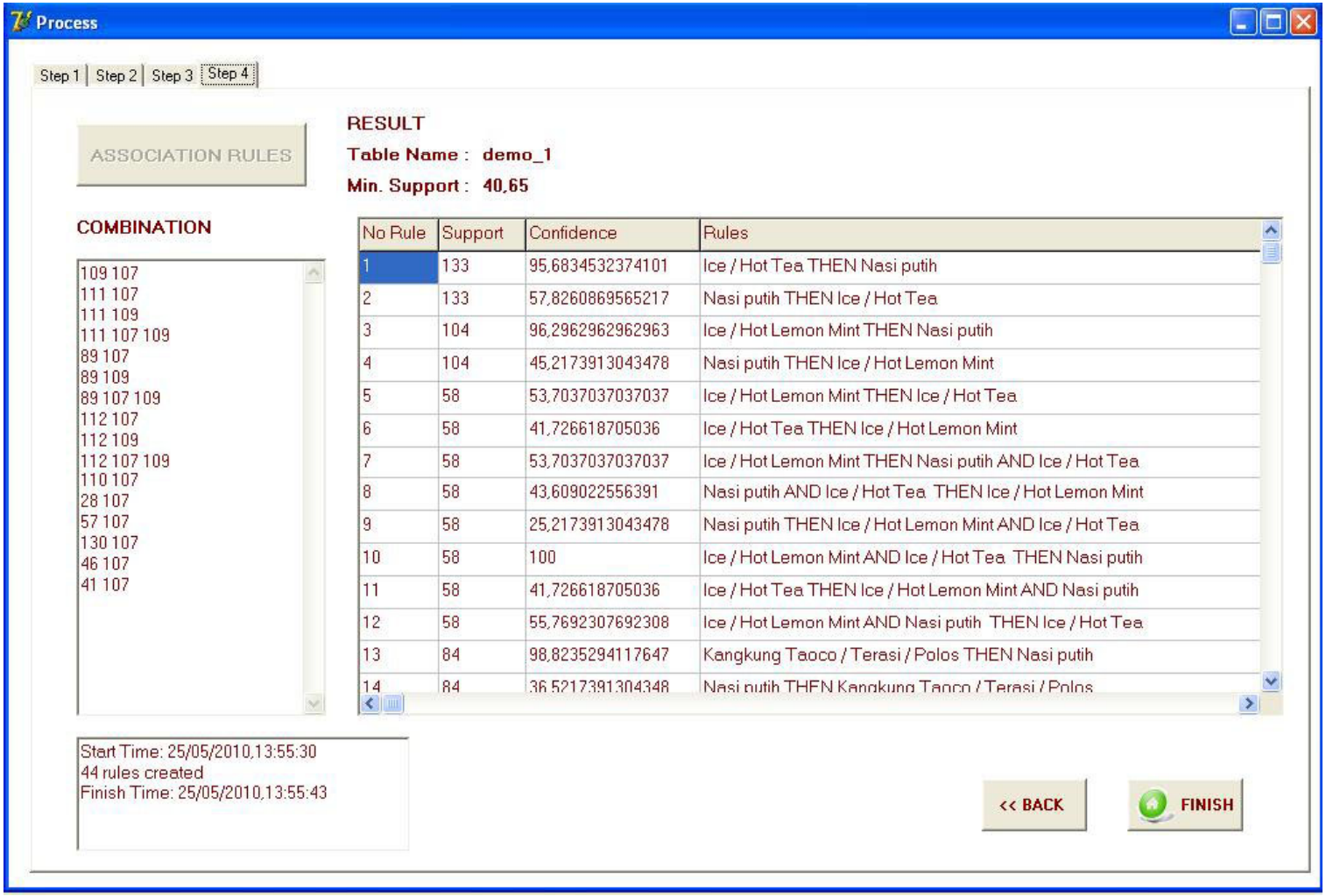
Create Tree


Process
Step 1
Step 2
Step 3
Step 4
TREE
{}
107 230
109 133
111 58
28 5
130 3
46 2
41 1
89 21
28 2
46 1
110 5
57 2
41 1
46 2
112 8
110 2
28 1
57 1
41 1
28 1
57 1
130 1
46 1
41 1
41 2
57 1
130 1
CREATE TREE
<< BACK
NEXT >>
Menu Name :
Udang Bago Bakar Madu
Start Time: 25/05/2010,13:55:25
Tree Created
Finish Time: 25/05/2010,13:55:28

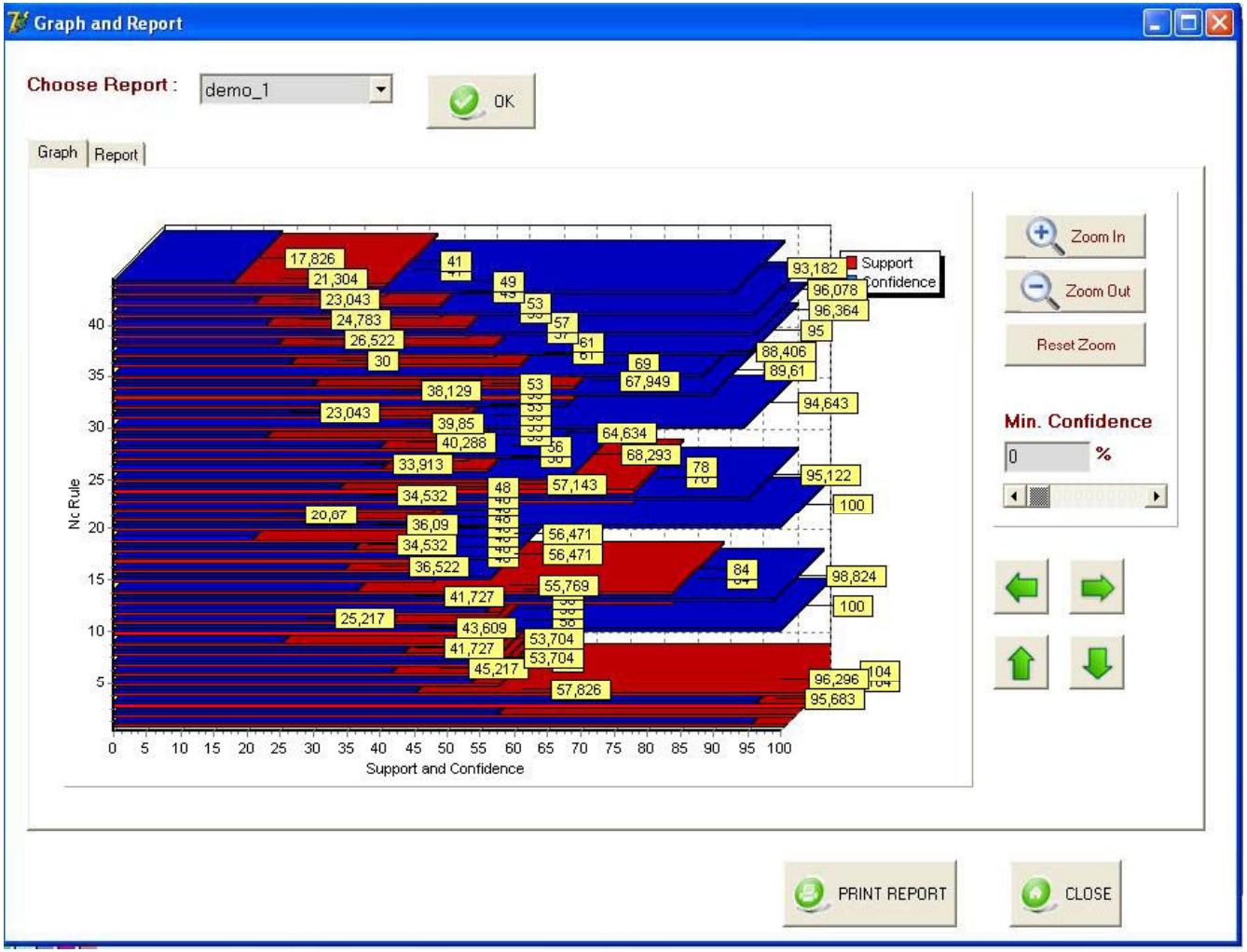
Conditional Pattern Base and Conditional FP-Tree



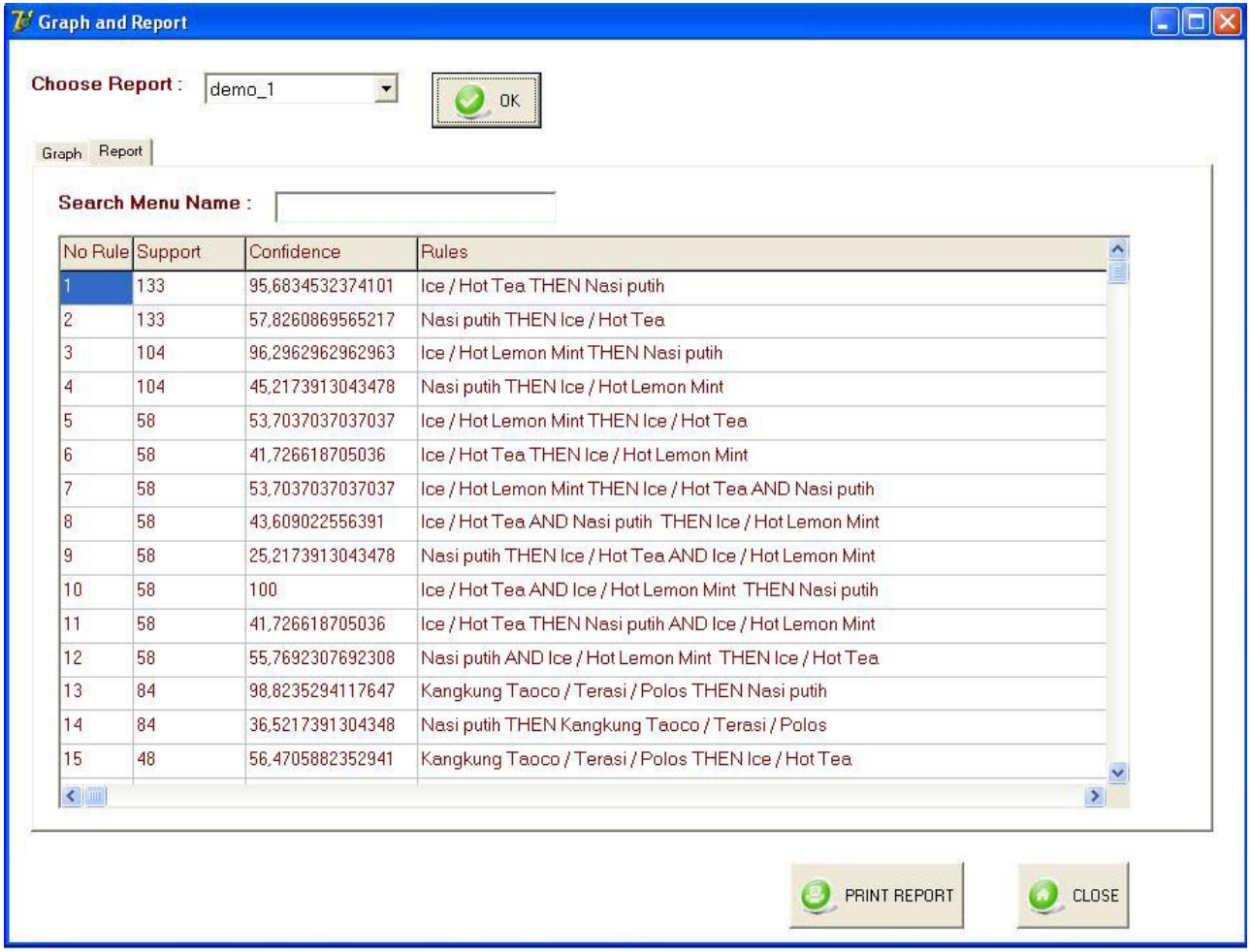
Generate Association Rules



Graph


Graph and Report
Choose Report :
demo_1
OK
Graph
Report
Zoom In
Zoom Out
Reset Zoom
Min. Confidence
0
%
Support
Confidence
Nc Rule
Support and Confidence
PRINT REPORT
CLOSE

Report



No Rule	Support	Confidence	Rules
1	133	95,6834532374101	Ice / Hot Tea THEN Nasi putih
2	133	57,8260869565217	Nasi putih THEN Ice / Hot Tea
3	104	96,2962962962963	Ice / Hot Lemon Mint THEN Nasi putih
4	104	45,2173913043478	Nasi putih THEN Ice / Hot Lemon Mint
5	58	53,7037037037037	Ice / Hot Lemon Mint THEN Ice / Hot Tea
6	58	41,726618705036	Ice / Hot Tea THEN Ice / Hot Lemon Mint
7	58	53,7037037037037	Ice / Hot Lemon Mint THEN Ice / Hot Tea AND Nasi putih
8	58	43,609022556391	Ice / Hot Tea AND Nasi putih THEN Ice / Hot Lemon Mint
9	58	25,2173913043478	Nasi putih THEN Ice / Hot Tea AND Ice / Hot Lemon Mint
10	58	100	Ice / Hot Tea AND Ice / Hot Lemon Mint THEN Nasi putih
11	58	41,726618705036	Ice / Hot Tea THEN Nasi putih AND Ice / Hot Lemon Mint
12	58	55,7692307692308	Nasi putih AND Ice / Hot Lemon Mint THEN Ice / Hot Tea
13	84	98,8235294117647	Kangkung Taoco / Terasi / Polos THEN Nasi putih
14	84	36,5217391304348	Nasi putih THEN Kangkung Taoco / Terasi / Polos
15	48	56,4705882352941	Kangkung Taoco / Terasi / Polos THEN Ice / Hot Tea

Preprocessing Testing Result

No	Transaction Period	Total Transaction	Total Menu	Process Time
1	1 day	66	101	3 s
2	3 days	178	133	10 s
3	1 week	271	143	16 s
4	2 week	403	152	24 s
5	1 month	827	163	54 s

Process Testing Result

No	Total Transaction's Notes	Min. Support	Total Rules	Process Time
1	66	10% (7)	274	1 m 15 s
		15% (10)	130	42 s
2	178	10% (18)	369	2 m 53 s
		15% (27)	58	49 s
3	271	10% (28)	292	3m 27 s
		15% (41)	44	1 m 2 s
4	403	10% (41)	260	4 m 48 s
		15% (61)	42	1 m 41 s
5	827	10% (83)	134	9 m 3 s
		15% (125)	26	3m 48 s

Questionnaire

No	Name	Position	Criteria				
			.	.	.	.	.
1	Eko	Owner	4	5	4	4	4
2	Handoko	General Manager	4	3	4	4	3
3	Sari	Cashier 1	4	4	3	3	4
4	Anik	Cashier 2	4	4	3	3	4
		Average :	4.0	4.0	3.5	3.5	3.75

Criteria Description :

1. The accuracy of the information generated.
2. Interface Design.
3. Ease of use of the program.
4. The use of language in the information.
5. Instructions given to the user in using application.

Conclusion

- If the specified minimum support getting smaller, then the generated frequent itemsets will become considerable, so the process time become longer.
- The result from mining process can displaying a correlation between data (association rules) with the support information and confidence that can be analyzed. This information will give additional consideration for user in further decision making.

Conclusion (continue)

- This application can display the result of rules into graph and table to view the result.
- Total time generated in the preprocessing and process with minimum support 10% is: one day transaction took 1 minute 18 seconds, three days transaction took 3 minutes 3 seconds, one week transaction took 3 minutes 43 seconds, two weeks transaction took 5 minutes 12 seconds, one month transaction took 9 minutes 57 seconds.

Conclusion (continue)

- Based on the experiments, this application is useable and can be implemented successfully which shown from questionnaire result from user as follows: level of information accuracy 80%, design interface 80%, simplicity in using the program 70%, language in providing information 70%, and guidance in using application 75%.

Suggestion

- With this method was expected to encourage the creation and implementation of better algorithms because there are many algorithms, like Pincer Search and Hash Based, which may make the analysis more faster and efficient in the processing time.

Suggestion (continue)

- Expected the interface design become more interesting and user friendly
- Applications can be developed into a web form, so users can use this application not only on a single computer and can be accessed wherever the user as long as they have the internet connection.

DESIGN AND MAKE OF DATA MINING MARKET BASKET ANALYSIS APLICATION AT DE JOGLO RESTAURANT

Nadia Andreani Dwiyono
26406019