

# **DESIGN AND MAKE OF DATA MINING MARKET BASKET ANALYSIS APPLICATION 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.

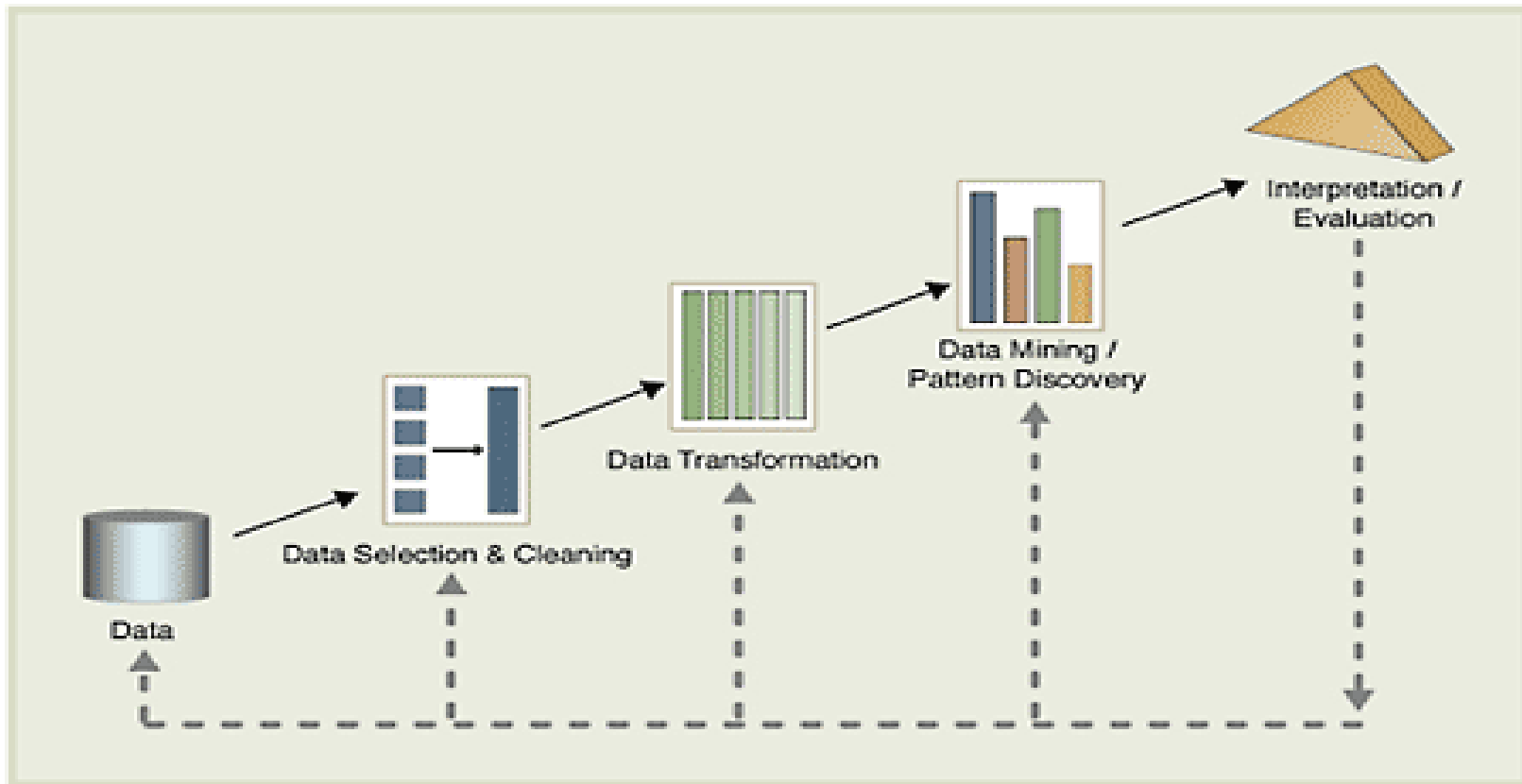


# Data Mining

- 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)





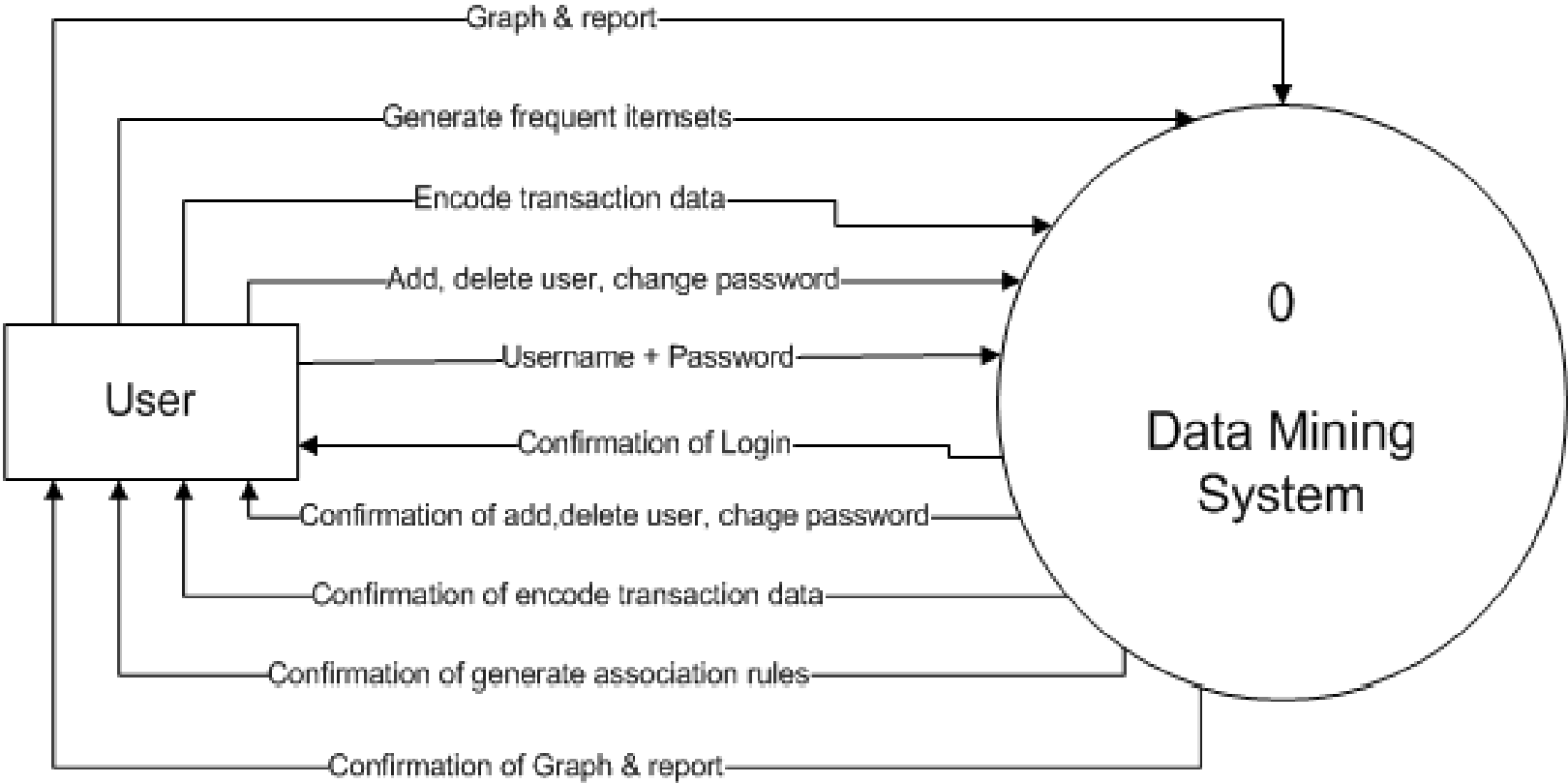
# Market Basket Analysis

- *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.

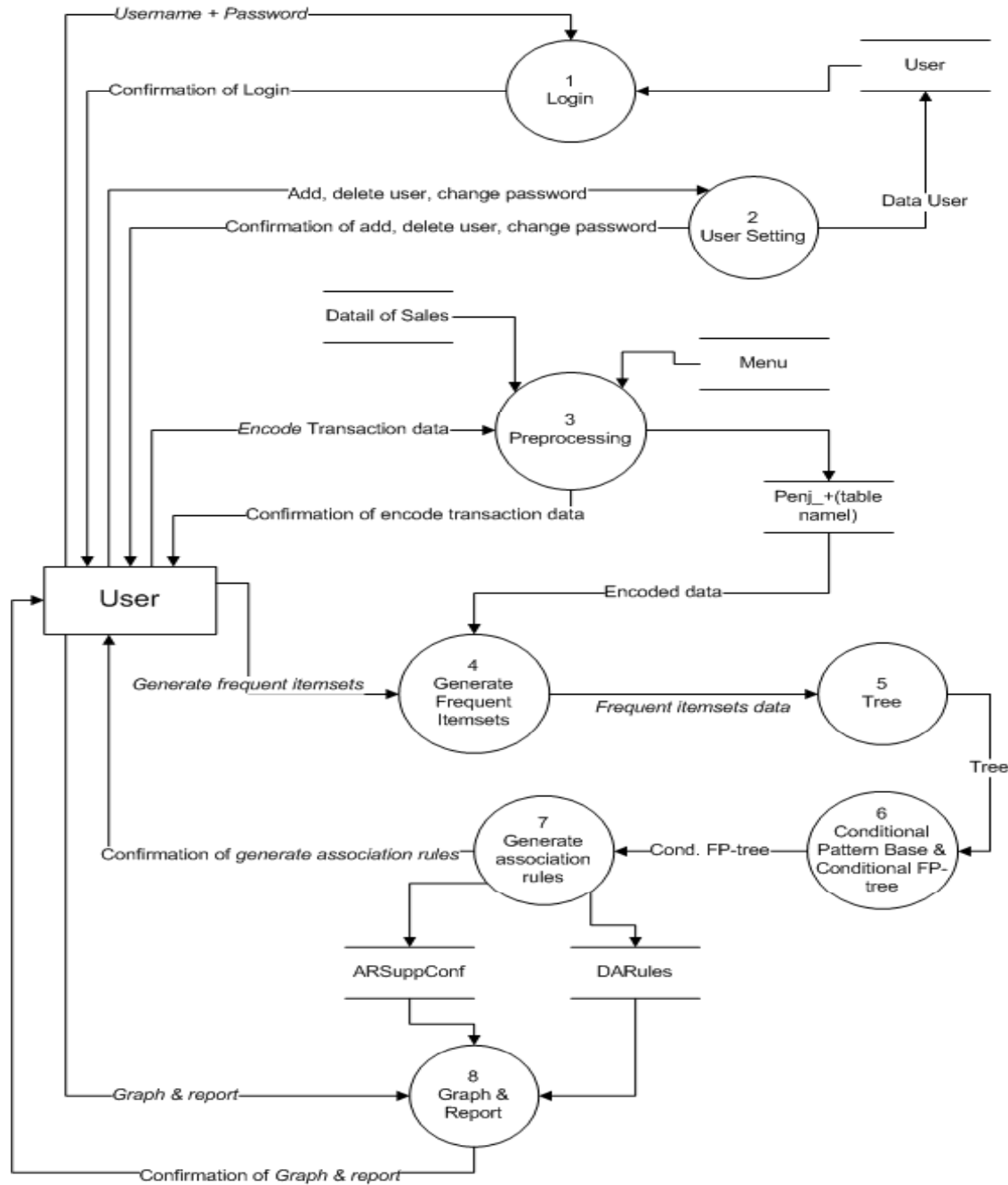
# FP-TREE

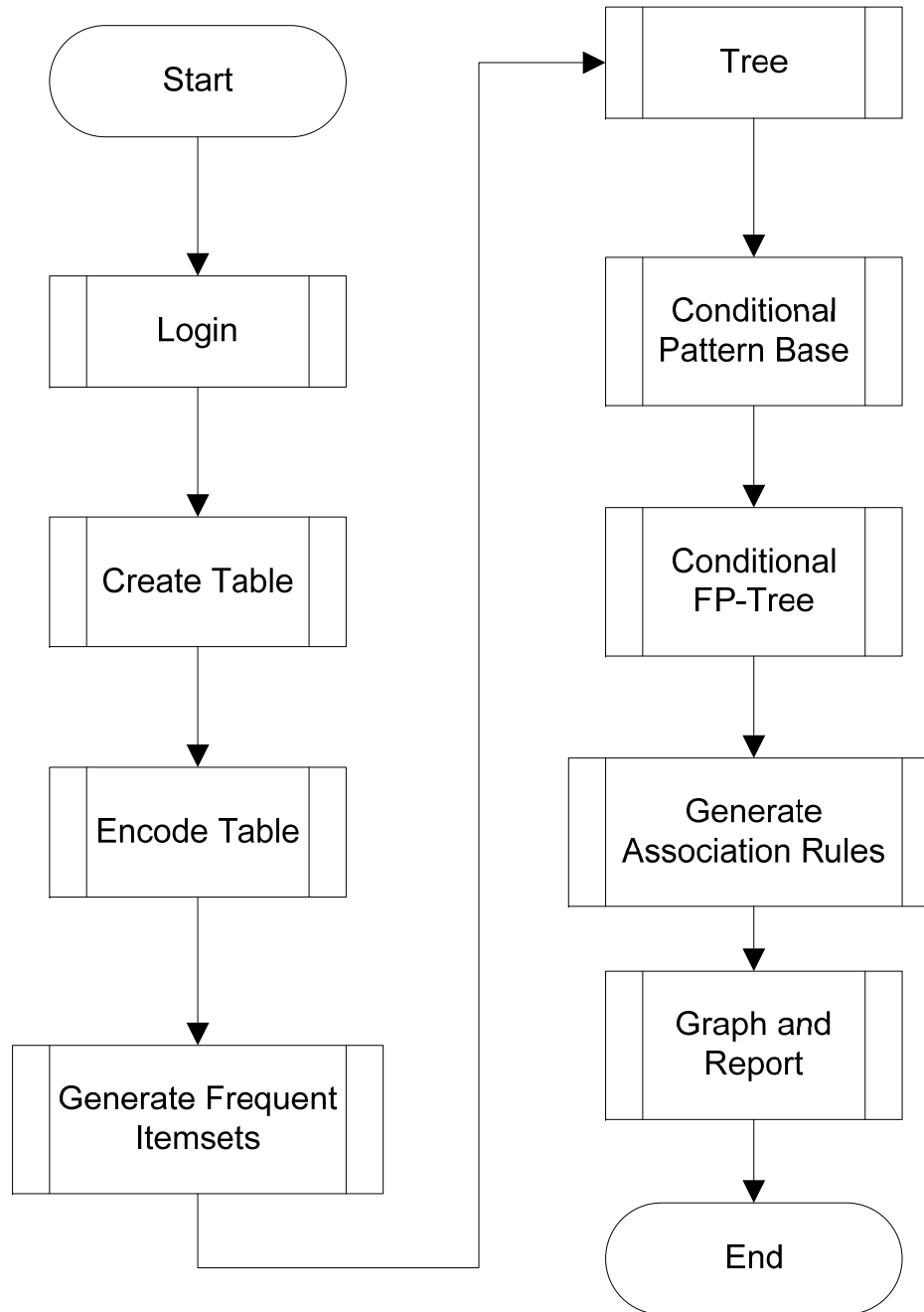
- FP-Tree is an algorithm used to find association patterns in transaction databases.
- Each node in the FP-Tree contains three important information. The informations are:
  - Label item
  - Count
  - Link

# Context Diagram



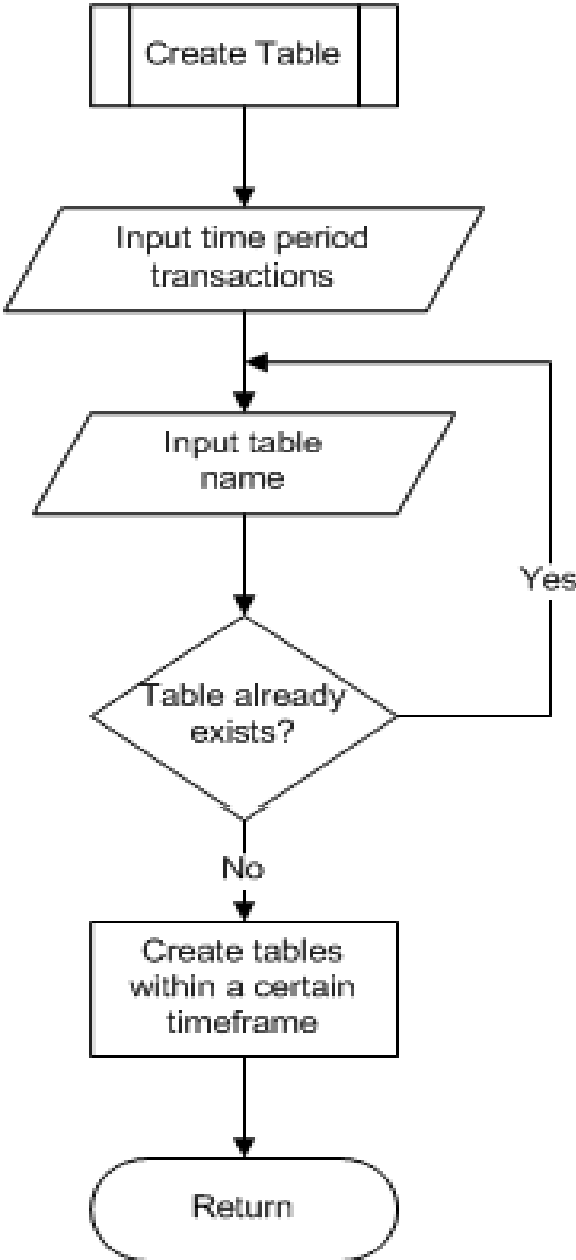
# DFD Level 0

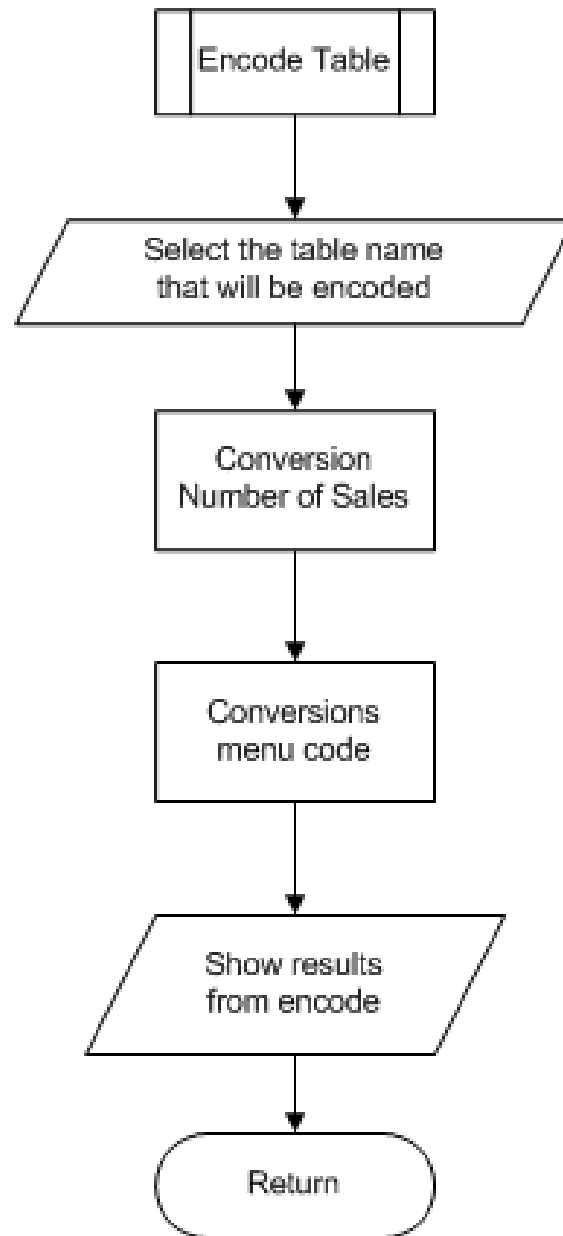
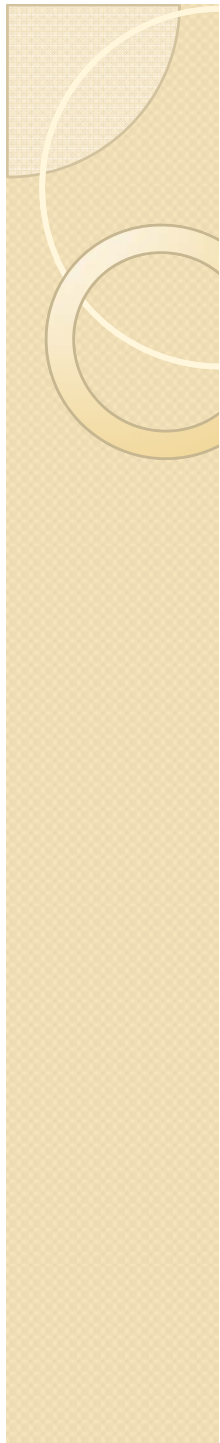




# Flowchart

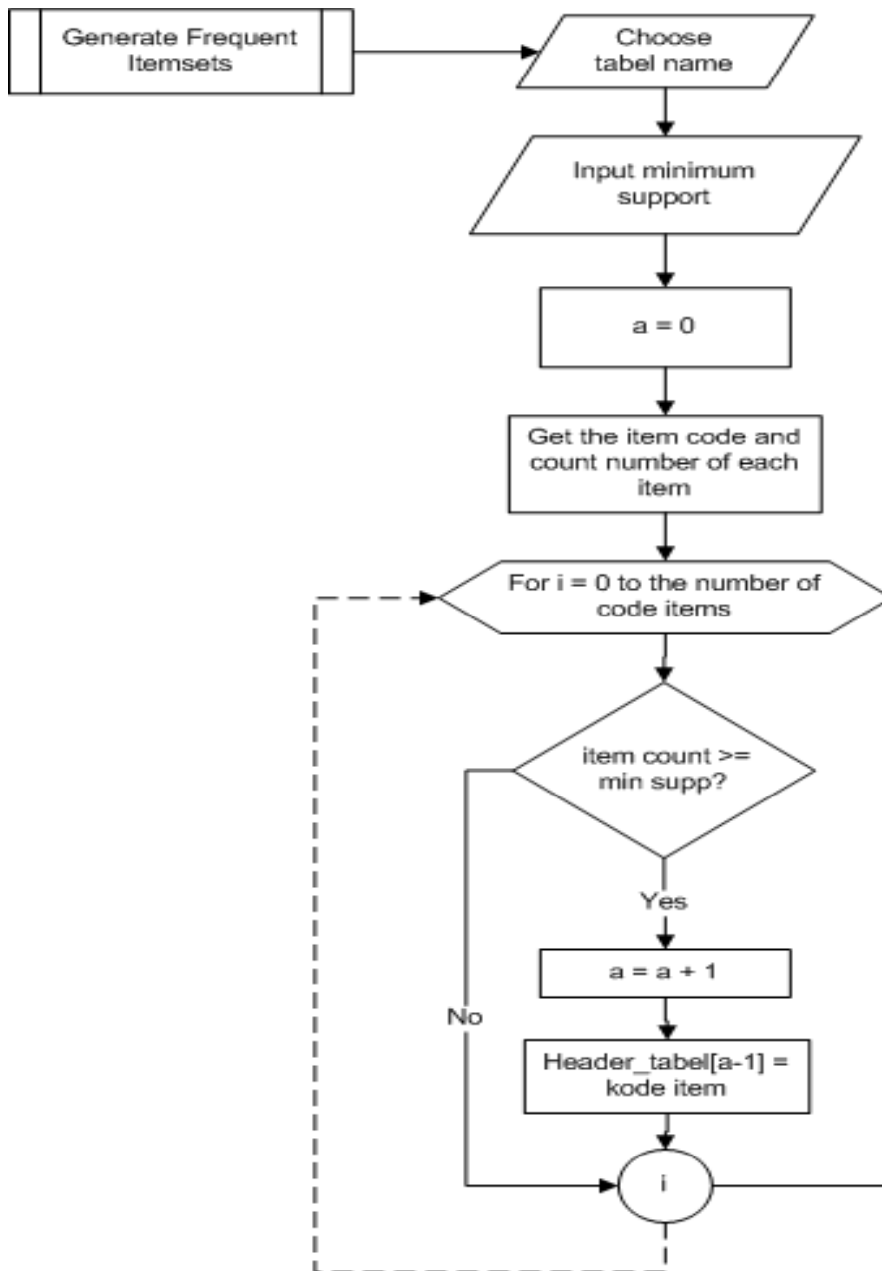
# Create Table Flowchart



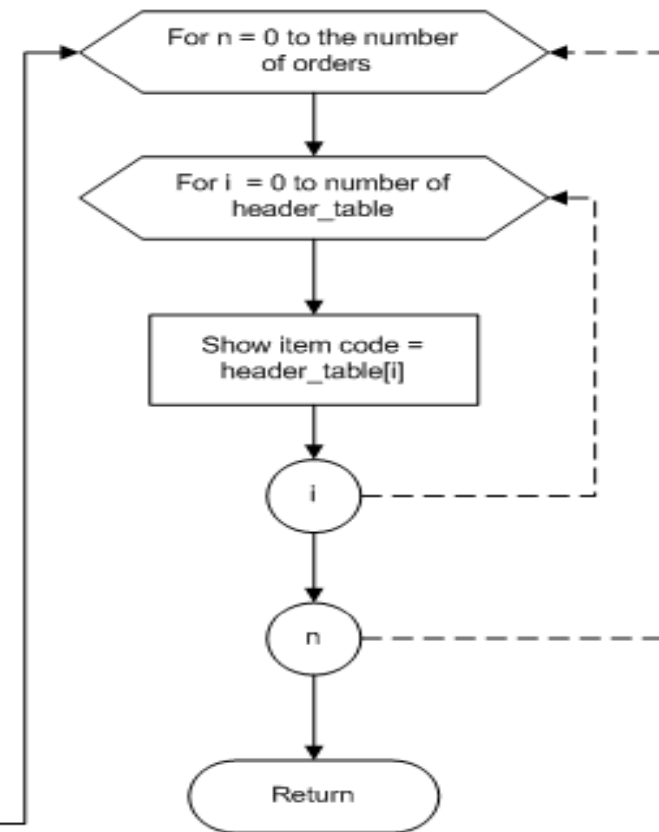


## Encode Table Flowchart

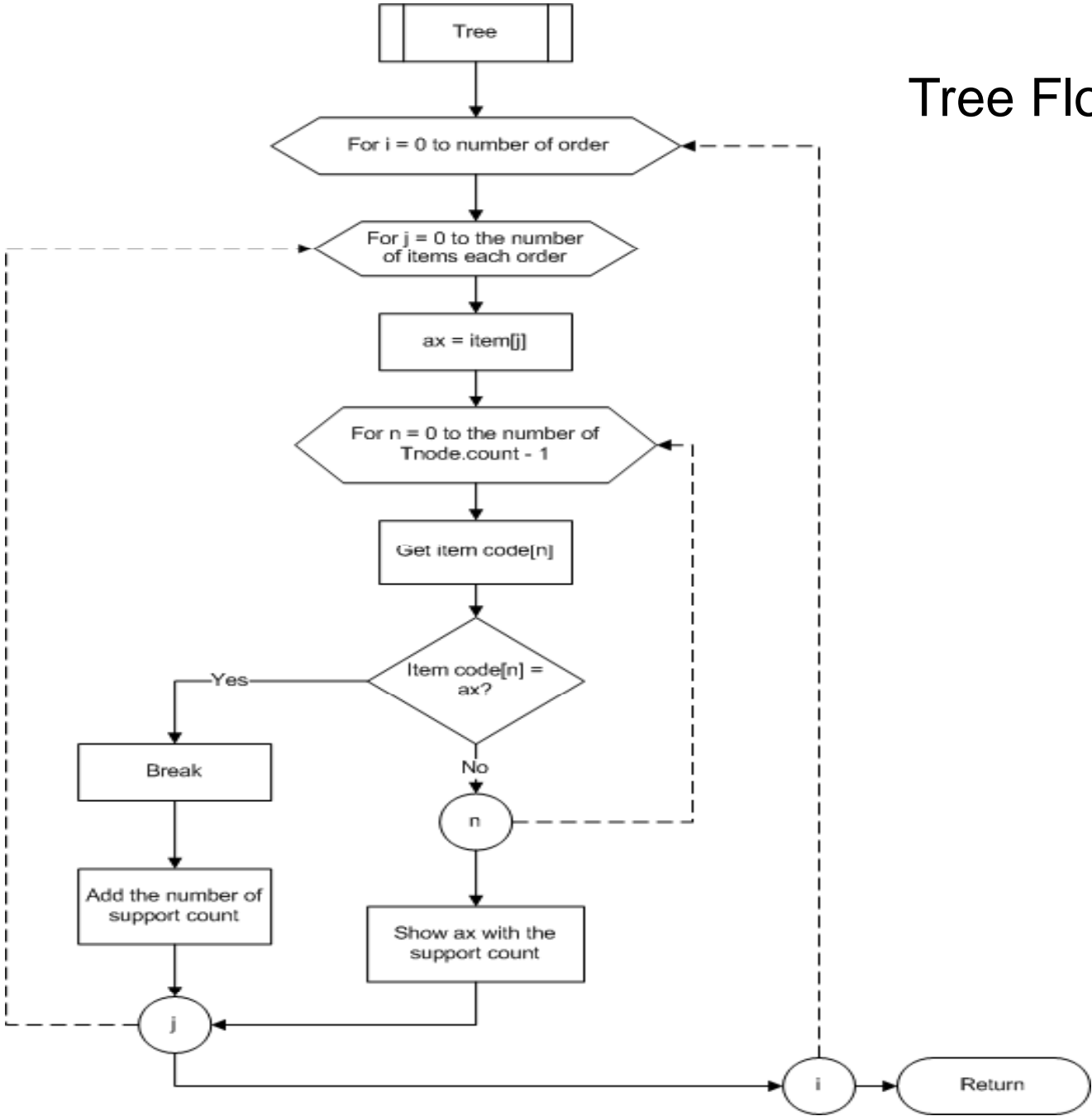




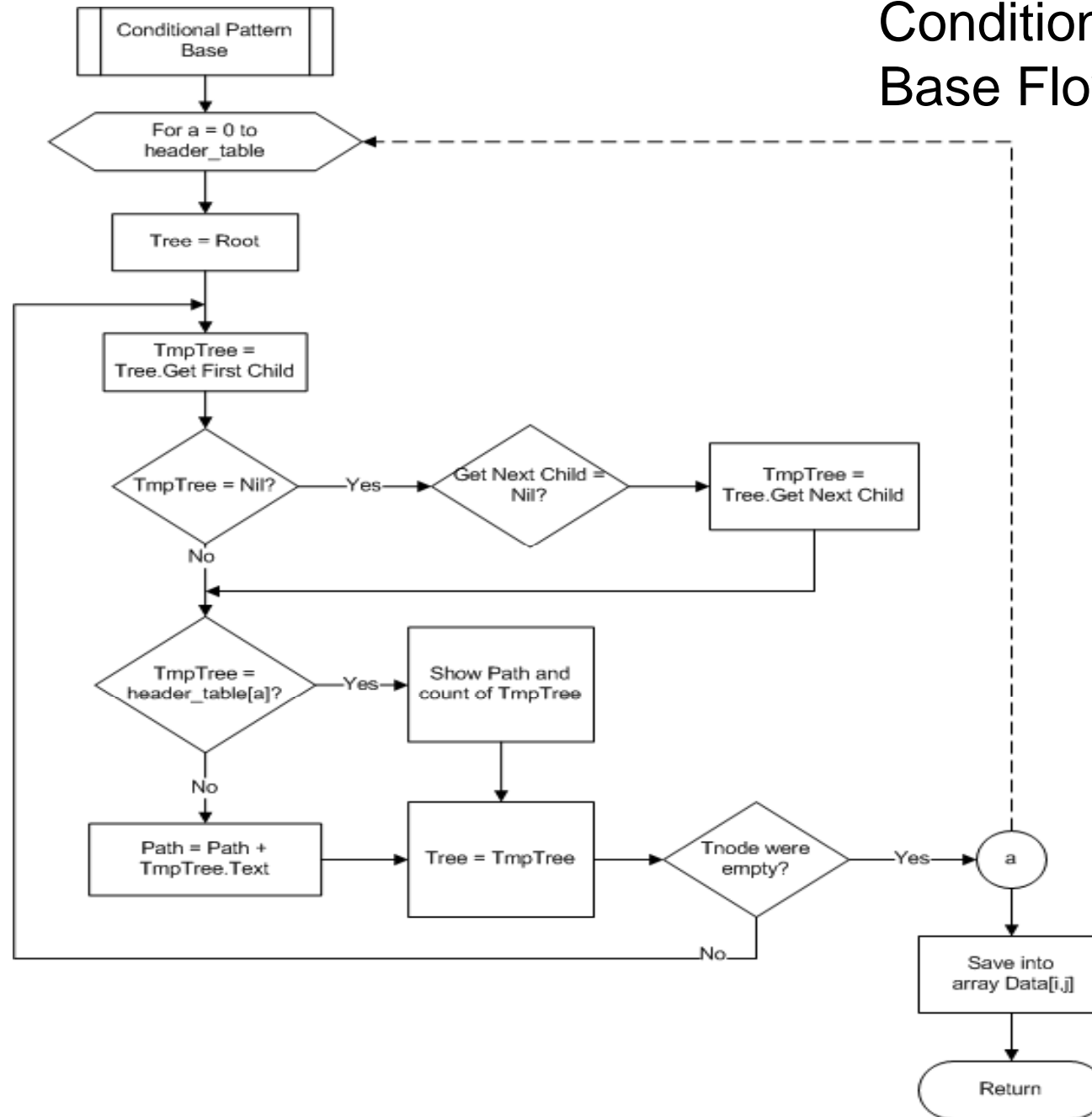
## Generate Frequent Item Flowchart



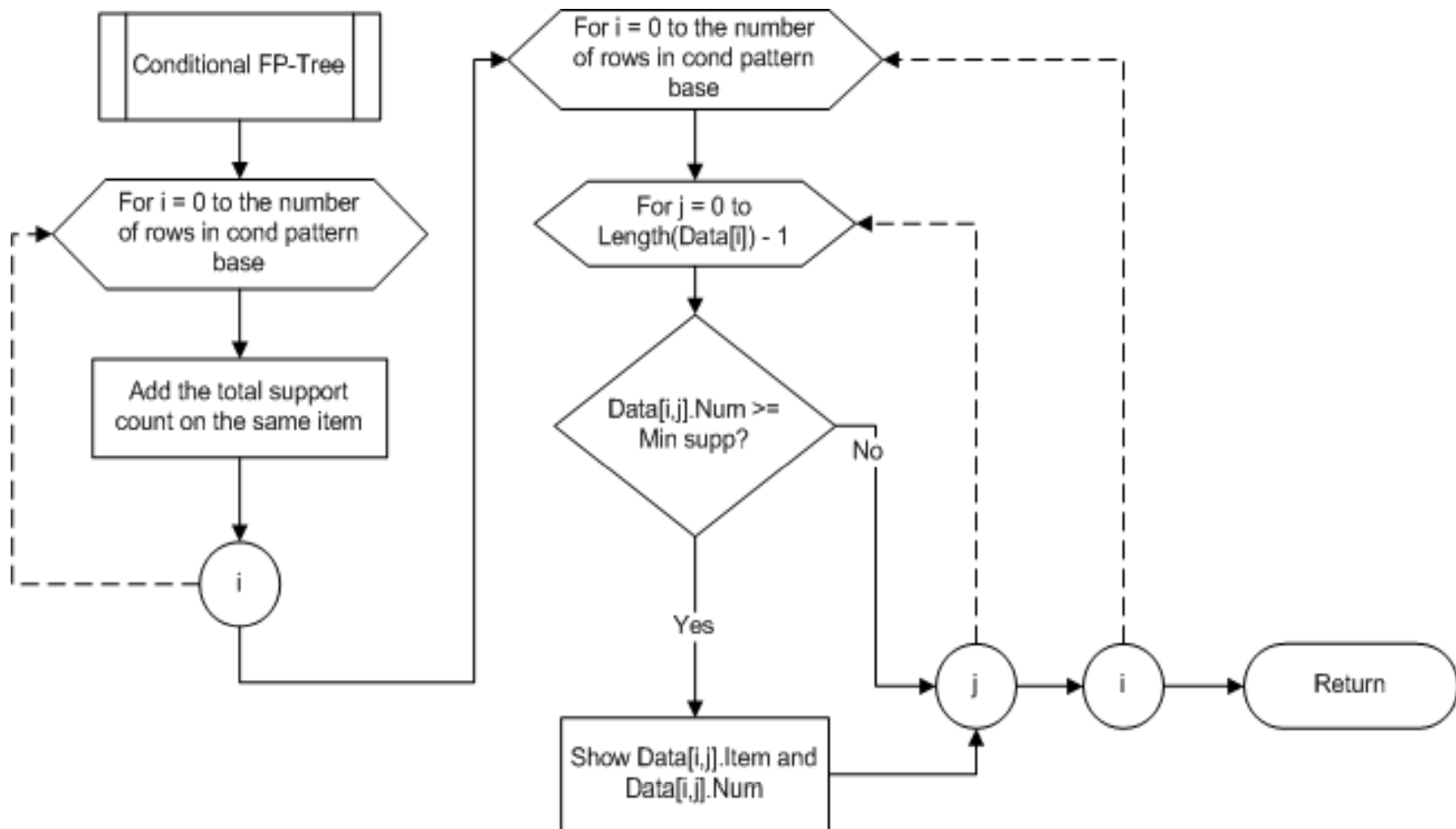
# Tree Flowchart



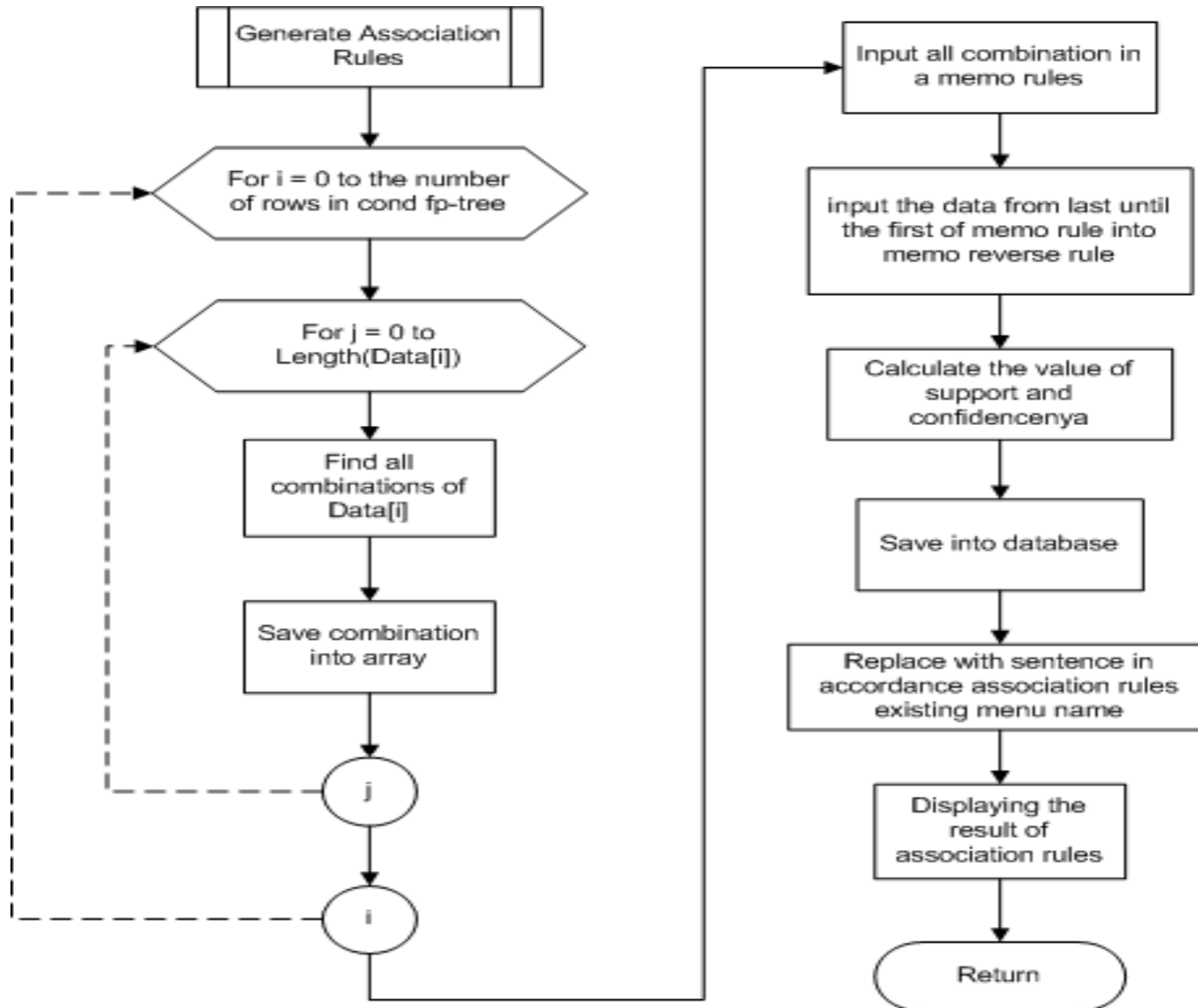
# Conditional Pattern Base Flowchart



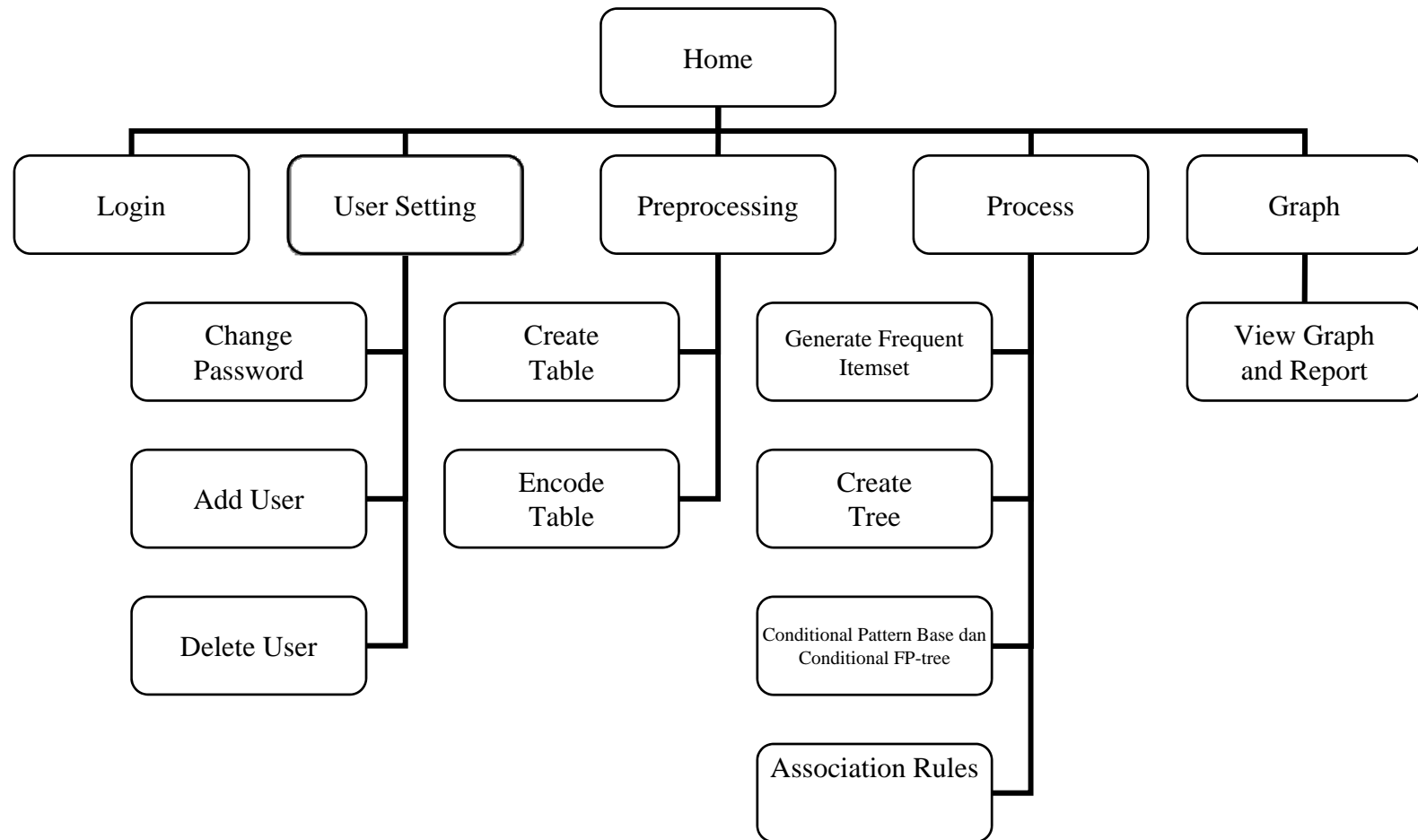
# Conditional FP-Tree Flowchart



# Generate Association Rules Flowchart



# System Design (Menu)



# Preprocessing (Create Table)

The screenshot shows a software window titled "Preprocessing" with a blue title bar. The window is divided into two steps: "Step 1" and "Step 2". The "Date" section contains two date pickers: "From" (01/10/2008) and "To" (05/10/2008), with an "ADD" button. A table to the right shows a date range from 10/01/2008 to 10/05/2008. The "Table Name" section has an input field containing "demo\_1" and a "CREATE" button. A "Success" dialog box is open, displaying "Table Created" and an "OK" button. A log box at the bottom left shows the start and finish times as 26/05/2010, 13:25:44. At the bottom of the window are "NEXT >>" and "CLOSE" buttons.

**Date**

**From** 01/10/2008

**To** 05/10/2008

From	To
10/01/2008	10/05/2008

**Table Name**

demo\_1

Start Time : 26/05/2010, 13:25:44  
Table Created  
Finish Time : 26/05/2010, 13:25:44


Success  
Table Created  
OK

NEXT >> CLOSE

# Preprocessing (Encode Table)

Preprocessing

Step 1 Step 2

Choose Table  

**Konversi Order**

No Nota	Konversi
20081001P-0001	1
20081001P-0002	2
20081001P-0003	3
20081001P-0004	4
20081001P-0005	5
20081001P-0006	6
20081001P-0007	7
20081001P-0008	8
20081001P-0009	9
20081001P-0010	10
20081001P-0011	11


**Konversi Menu**

Menu	Konversi
1011	1
1012	2
1013	3
	4
	5
	6
	7
1019	8
1020	9
1021	10
1030	11





Start Time:  
26/05/2010,13:26:29  
271 Order ID Rows  
Affected  
143 Menu ID Rows  
Affected  
Finish Time:  
26/05/2010,13:26:44

<< BACK 




# Generate Frequent Item

Process

Step 1 | Step 2 | Step 3 | Step 4

Choose Table: demo\_1

Min. Support: 15 % OR 40.65 



Max Support: 271

No Menu	Total
107	230
109	139
111	108
89	85
112	82
110	77
28	69
57	60
130	55
46	51
41	44

No Order	No Menu	Order Date
1	107	01/10/2008
1	109	01/10/2008
1	111	01/10/2008
2	107	01/10/2008
2	109	01/10/2008
3	107	01/10/2008
3	109	01/10/2008
3	111	01/10/2008
4	107	01/10/2008
4	111	01/10/2008
4	89	01/10/2008
4	46	01/10/2008

107 109 111  
107 109  
107 109 111  
107 111 89 46  
107 110  
107 111 112 130  
107 111  
107 110 28 130  
107 111 110  
107 109 111 28 130  
107 109 111 89 28  
107 109 111 89 28 46  
107 109 111 110 28 46  
107 111 89 112 110 28  
107 109 112 110 28 57  
107 111  
107 109 111 41  
107 109 89 28 57 46  
107 111 89  
107 109 28  
107 110 130 46  
107 109 89 112 110 57 130  
107 109 111 89 110 57 41

Start Time: 25/05/2010,13:53:39  
Transaction Notes: 271  
Finish Time: 25/05/2010,13:54:39

# Create Tree

The screenshot shows a software window titled "Process" with a blue title bar. At the top, there are four tabs labeled "Step 1", "Step 2", "Step 3", and "Step 4". The main area is divided into two sections. On the left, under the heading "TREE", is a hierarchical tree structure. The root node is a folder icon containing a brace symbol. Below it are several levels of nodes, each represented by a folder icon and a number. The nodes are: 107 230, 109 133, 111 58, 28 5, 130 3, 46 2, 41 1, 89 21, 28 2, 46 1, 110 5, 57 2 (highlighted in blue), 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, and 130 1. On the right side, there are three buttons: "CREATE TREE" (top), "<< BACK" (middle), and "NEXT >>" (middle). Below these buttons is a text box containing the text "Menu Name : Udang Bago Bakar Madu". At the bottom right, there is a status box with the following text: "Start Time: 25/05/2010,13:55:25", "Tree Created", and "Finish Time: 25/05/2010,13:55:28".

# Conditional Pattern Base and Conditional FP-Tree

Step 1 | Step 2 | **Step 3** | Step 4

### CONDITIONAL PATTERN BASE

```
107 :  
109 : 107-133 |  
111 : 107 109-58 | 107-46 |  
89 : 107 109 111-21 | 107 109-27 | 107 111-15 | 107-21 |  
112 : 107 109 111 89-8 | 107 109 111-11 | 107 109-23 | 107 109 89-11 | 107 111 89-6 | 107 111-4 | 107-7 | 107  
110 : 107 109 111 89-5 | 107 109 111 89 112-2 | 107 109 111-8 | 107 109 111 112-5 | 107 109 112-8 | 107 109  
28 : 107 109 111-5 | 107 109 111 89-2 | 107 109 111 89 112 110-1 | 107 109 111 89 112-1 | 107 109 111 110-  
57 : 107 109 111 89 110-2 | 107 109 111 89 112 110 28-1 | 107 109 111 89 112 28-1 | 107 109 111 89 112-1 |  
130 : 107 109 111 28-3 | 107 109 111 89 112 28 57-1 | 107 109 111 89 112 57-1 | 107 109 111 89 112-1 | 107  
46 : 107 109 111 28-2 | 107 109 111 89 28-1 | 107 109 111 89-2 | 107 109 111 89 112 28 57 130-1 | 107 109 1  
41 : 107 109 111 28 46-1 | 107 109 111 89 110 57-1 | 107 109 111 89 112 110-1 | 107 109 111 89 112 28 57 1
```

**CREATE CONDITIONAL**

**<< BACK**      **NEXT >>**

### CONDITIONAL FP TREE

```
107:  
109:107-133 |  
111:107-104 | 109-58 |  
89:107-84 | 109-48 |  
112:107-78 | 109-56 |  
110:107-69 |  
28:107-61 |  
57:107-57 |  
130:107-53 |  
46:107-49 |  
41:107-41 |
```

Start Time: 25/05/2010,13:55:29  
Finish Time: 25/05/2010,13:55:29

# Generate Association Rules

7 Process

Step 1 | Step 2 | Step 3 | **Step 4**

ASSOCIATION RULES

COMBINATION

109 107  
111 107  
111 109  
111 107 109  
89 107  
89 109  
89 107 109  
112 107  
112 109  
112 107 109  
110 107  
28 107  
57 107  
130 107  
46 107  
41 107

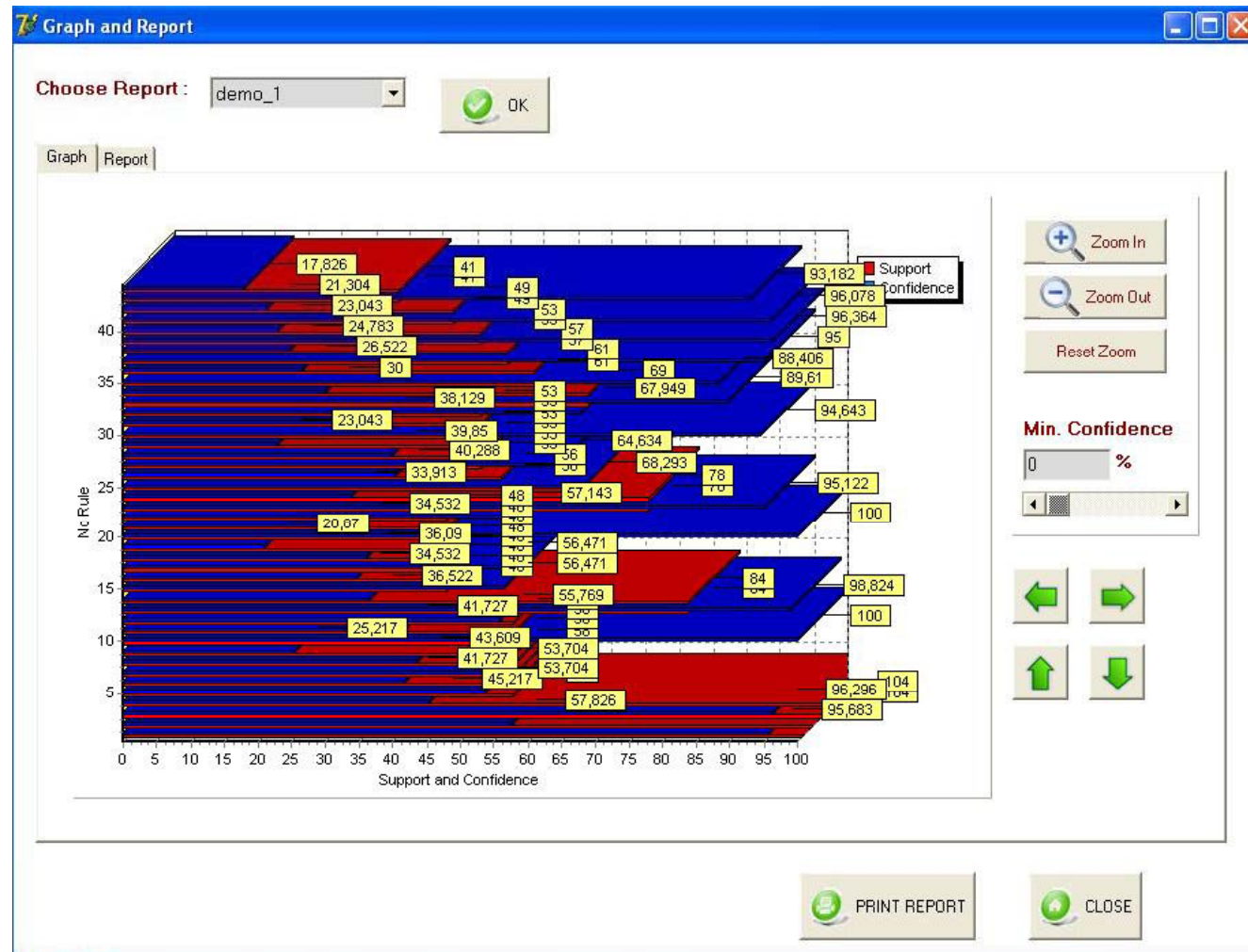
**RESULT**  
Table Name : demo\_1  
Min. Support : 40,65

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 Nasi putih AND Ice / Hot Tea
8	58	43,609022556391	Nasi putih AND Ice / Hot Tea THEN Ice / Hot Lemon Mint
9	58	25,2173913043478	Nasi putih THEN Ice / Hot Lemon Mint AND Ice / Hot Tea
10	58	100	Ice / Hot Lemon Mint AND Ice / Hot Tea THEN Nasi putih
11	58	41,726618705036	Ice / Hot Tea THEN Ice / Hot Lemon Mint AND Nasi putih
12	58	55,7692307692308	Ice / Hot Lemon Mint AND Nasi putih 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

Start Time: 25/05/2010,13:55:30  
44 rules created  
Finish Time: 25/05/2010,13:55:43


<< BACK FINISH

# Graph



# Report



Graph and Report

Choose Report : demo\_1 

Graph Report

Search Menu Name :

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 APPLICATION AT DE JOGLO RESTAURANT**



**Nadia Andreani Dwiyono  
26406019**