

## BAB IV

### ANALISIS DATA

#### 4.1 DATA

Data ini adalah data yang diambil di sebuah minimarket di daerah Perumnas Klender Jakarta Timur pada 6 Juni 2008. Data tersebut berisi nama jenis *item* yang dibeli pada tiap transaksi. Terdapat 100 transaksi dan 50 jenis *item*. Untuk lebih jelasnya, data yang diambil dapat dilihat pada Lampiran 1. Data tersebut dianalisis dengan menggunakan metode kaidah asosiasi dengan bantuan perangkat lunak SPSS 14, Microsoft Office Excel.

#### 4.2 ASUMSI

Berikut ini adalah asumsi-asumsi yang dipakai dalam penelitian ini:

1. Semua minimarket menjual *item-item* yang diteliti
2. Penulisan data berdasarkan struk yang keluar di kasir. Siapa yang membeli dan kapan membelinya diabaikan.
3. *Item* yang dimaksud adalah jenis produk dengan mengabaikan merek.

### 4.3 PERMASALAHAN PENELITIAN

Permasalahan yang terjadi pada penelitian ini adalah bagaimana menemukan aturan-aturan asosiasi pada data tersebut.

### 4.4 LANGKAH-LANGKAH

Berikut ini adalah langkah-langkah dalam proses pencarian kombinasi *item* yang sering dibeli oleh pelanggan di minimarket.

#### 4.4.1 Transformasi Data Dalam Bentuk Data Transaksi

Untuk mentransformasi data tersebut ke bentuk data transaksi, pertama harus didefinisikan variabel-variabel yang menggambarkan jenis *item* yang terjual pada data tersebut. Misalkan ada himpunan  $\{X_1, X_2, X_3, \dots, X_{50}\}$ , di mana himpunan tersebut adalah himpunan peubah acak yang menggambarkan *item* yang dijual oleh salah satu minimarket di Perumnas Klender. Tabel 4.1 adalah definisi dari variabel-variabel yang digunakan dalam penelitian ini:

**Tabel 4.1** Definisi Masing-masing Peubah Acak yang Akan Digunakan dalam Pemrosesan Metode Kaidah Asosiasi

$X_1$	Air Mineral	$X_{11}$	Kaos kaki	$X_{21}$	Garam	$X_{31}$	Sikat gigi	$X_{41}$	Korek
$X_2$	Soft Drink	$X_{12}$	Saus	$X_{22}$	Gula	$X_{32}$	Shampoo	$X_{42}$	Rokok
$X_3$	Kopi	$X_{13}$	Telur	$X_{23}$	Minyak goreng	$X_{33}$	Pelembab	$X_{43}$	Tissue
$X_4$	Teh	$X_{14}$	Gula	$X_{24}$	Obat	$X_{34}$	Pengharum baju	$X_{44}$	Bedak
$X_5$	Susu	$X_{15}$	Kue kering	$X_{25}$	Agar-agar	$X_{35}$	Penyegar mulut	$X_{45}$	Parfum
$X_6$	Susu Bayi	$X_{16}$	Chiki	$X_{26}$	Pembalut	$X_{36}$	Pengharum ruangan	$X_{46}$	Batre
$X_7$	roti	$X_{17}$	Coklat	$X_{27}$	Popok	$X_{37}$	Pembersih lantai	$X_{47}$	Majalah
$X_8$	Keju	$X_{18}$	Ice cream	$X_{28}$	Sabun mandi	$X_{38}$	m. kayu putih	$X_{48}$	Tepung
$X_9$	Mentega	$X_{19}$	Permen	$X_{29}$	Sabun cuci	$X_{39}$	Obat nyamuk	$X_{49}$	Sandal
$X_{10}$	Mi instan	$X_{20}$	Sarden	$X_{30}$	Pasta gigi	$X_{40}$	Kondom	$X_{50}$	Sirup

Dengan memasukkan nilai dari  $X_j$  yaitu

$X_j : x_{ij} = 1$ , jika *item* ke-j dibeli pada transaksi ke-i  
 $x_{ij} = 0$ , jika *item* ke-j tidak dibeli pada transaksi ke-i

$i = 1, 2, \dots, 100$  dan  $j = 1, 2, \dots, 50$  data tersebut dapat ditransformasi ke bentuk data transaksi. Untuk lebih jelasnya format data transaksi dari data tersebut, dapat dilihat pada Lampiran 2.

#### 4.4.2 Pencarian Aturan-Aturan Asosiasi Antar *Item*

Dalam pemrosesan metode kaidah asosiasi, ada tahap untuk mencari aturan-aturan asosiasi antar *item*. Aturan-aturan tersebut nantinya akan memberikan informasi mengenai kombinasi antar *item* yang sering dibeli oleh pelanggan suatu minimarket. Pencarian aturan-aturan asosiasi pada tugas akhir ini dilakukan dengan menggunakan algoritma Apriori.

Algoritma Apriori dilakukan dengan dua tahap, yaitu pencarian *itemset-itemset* yang *frequent* dan pembentukan aturan asosiasi dari *itemset-itemset* yang *frequent* tersebut. Untuk melakukan tahap-tahap tersebut, diberikan nilai *minimum support* dan *minimum confidence*. Pada penelitian ini *minimum support* = 0.25 dan *minimum confidence* = 0.65 yang ditentukan oleh penulis dan  $X = \{X_1, X_2, \dots, X_{50}\}$  adalah himpunan seluruh *item* yang dijual oleh minimarket tersebut.

##### 4.4.2.1 Tahap Pencarian *Itemset-Itemset* yang *Frequent*

Berikut ini adalah langkah-langkah dalam pencarian semua *itemset* yang *frequent*.

### a. Pencarian Frequent 1-itemset ( $F_1$ )

Scan data transaksi untuk mendapatkan *support* dari tiap 1-itemset (Properti Apriori 1). Bandingkan dengan *minimum support* untuk mendapatkan *frequent 1-itemset* ( $F_1$ ), hal ini sesuai dengan Properti Apriori 2.

Berdasarkan *output* pada lampiran 3, didapat nilai *itemset frequency* dari tiap 1-itemset dan berdasarkan persamaan 2, didapat nilai *support* dari  $X_j$  adalah:

$$S(X_j) = \frac{1}{N} \sum_{i=1}^N x_{ij}$$

$$= \frac{1}{100} \sum_{i=1}^{100} x_{ij} \quad , \quad j = 1, 2, \dots, 100.$$

Bandinkan nilai dari tiap  $S(X_j)$  dengan *minimum support* = 0,25, maka didapat himpunan

$$F_1 = \{X_j \mid S(X_j) = \frac{1}{100} \sum_{i=1}^{100} x_{ij} \geq \text{minimum support} = 0.25\}$$

Berdasarkan *output* di Lampiran 4, nilai-nilai *support* pada tiap 1-itemset dapat dilihat pada Tabel 4.2.

**Tabel 4.2.** 1-itemset dan Supportnya

itemset	support	itemset	support	itemset	support	itemset	support
X1	0,11	X14	0,09	X27	0,04	X40	0,04
X2	0,42	X15	0,24	X28	0,46	X41	0,1
X3	0,14	X16	0,19	X29	0,41	X42	0,1
X4	0,03	X17	0,37	X30	0,38	X43	0,13
X5	0,59	X18	0,09	X31	0,36	X44	0,02
X6	0,07	X19	0,16	X32	0,21	X45	0,17
X7	0,32	X20	0,03	X33	0,09	X46	0,01
X8	0,16	X21	0,01	X34	0,09	X47	0,03
X9	0,09	X22	0,03	X35	0,02	X48	0,01
X10	0,4	X23	0,25	X36	0,11	X49	0,04
X11	0,02	X24	0,2	X37	0,05	X50	0,02
X12	0,24	X25	0,01	X38	0,04		
X13	0,34	X26	0,14	X39	0,15		

Berdasarkan Tabel 4.2. didapat himpunan *frequent 1-itemset* ( $F_1$ ) yaitu:

$$F_1 = \{X_2, X_5, X_7, X_{10}, X_{13}, X_{17}, X_{28}, X_{29}, X_{30}, X_{31}\}.$$

**b. Pencarian *Frequent 2-itemset*. ( $F_2$ )**

Gunakan  $F_1$  join  $F_1$  untuk mendapatkan himpunan kandidat dari *frequent 2-itemset* ( $L_2$ ). Lalu scan database transaksi untuk mendapatkan *support* dari tiap kandidat *frequent 2-itemset* (Properti Apriori 1). Bandingkan dengan *minimum support* untuk mendapatkan *frequent 2-itemset* ( $F_2$ ) (Properti Apriori 2). Himpunan kandidat dari *frequent 2-itemset* adalah :

$$L_2 = \{X_p, X_q \mid X_p \in F_1, X_q \in F_1, p \neq q\}.$$

Berdasarkan *output* pada lampiran 4, didapat nilai *itemset frequency* dari tiap *2-itemset*. Dari nilai *itemset frequency*nya didapat nilai *support* dari tiap *2-itemset* sebagai berikut :

Misalkan  $A$  adalah *itemset* ,  $A \subseteq X$  dan  $\lambda = \{\text{indeks pada } A\}$  , *support*

$$\text{dari itemset } A \text{ adalah } S(A) = \frac{\text{Itemset frequency}(A)}{\text{total transaksi}} = \frac{1}{100} \sum_{i=1}^{100} \prod_{a \in \lambda} x_{ia}$$

Berdasarkan *output* pada Lampiran 5, nilai-nilai *support* dari tiap kandidat *frequent 2-itemset* dapat dilihat pada tabel 4.3.

**Tabel 4.3.** Kandidat *Frequent 2-itemset* dan *Support*-nya

<i>itemset</i>	Support	<i>itemset</i>	support	<i>itemset</i>	support	<i>itemset</i>	Support
X2X5	0,23	X5X28	0,26	X10X28	0,21	X28X29	0,31
X2X7	0,11	X5X29	0,26	X10X29	0,16	X28X30	0,21
X2X10	0,19	X5X30	0,19	X10X30	0,12	X28X31	0,2
X2X13	0,13	X5X31	0,15	X10X31	0,29	X29X30	0,18
X2X17	0,2	X7X10	0,12	X13X17	0,11	X29X31	0,19
X2X28	0,17	X7X13	0,15	X13X28	0,19	X30X31	0,29
X2X29	0,13	X7X17	0,13	X13X29	0,15		
X2X30	0,13	X7X28	0,12	X13X30	0,12		
X2X31	0,17	X7X29	0,15	X13X31	0,09		
X5X7	0,24	X7X30	0,13	X17X28	0,18		
X5X10	0,22	X7X31	0,11	X17X29	0,13		
X5X13	0,18	X10X13	0,26	X17X30	0,12		
X5X17	0,25	X10X17	0,13	X17X31	0,11		

Berdasarkan Tabel 4.3 , didapat himpunan *frequent 2-itemset* yaitu:

$$F_2 = \{\{X_{10}, X_{13}\}, \{X_{10}, X_{31}\}, \{X_{28}, X_{29}\}, \{X_{30}, X_{31}\}\}$$

### c. Pencarian *Frequent 3-itemset*. ( $F_3$ )

Gunakan  $F_2 \text{ join } F_2$  untuk mendapatkan himpunan kandidat dari *frequent 3-itemset*. Himpunan kandidat dari *frequent 3-itemset* dapat dilihat pada Tabel 4.4

**Tabel 4.4** Kandidat 3-*itemset*

Kandidat 3- <i>itemset</i>
X10X13X31
X10X30X31

Dapat dilihat kandidat yang terbentuk sebanyak dua kandidat, yaitu  $X_{10}, X_{13}, X_{31}$  (kejadian *item* mi instant, telur dan sikat gigi terbeli bersamaan) dan  $X_{10}, X_{30}, X_{31}$  (kejadian *item* mi instant, pasta gigi dan sikat gigi terbeli bersamaan). Gunakan properti Apriori 3 untuk memangkas kandidat yang memiliki *subset* yang tidak *frequent*. Karena semua kandidat dari *frequent 3-itemset* memiliki *subset* yang tidak *frequent*, maka tidak ada kandidat yang tersisa atau himpunan kandidat dari *frequent 3-itemset* =  $\emptyset$

#### 4.4.2.2 Tahap Pembentukan Aturan-Aturan Asosiasi

Prosedur pembentukan aturan-aturan asosiasi :

1. Untuk setiap *frequent itemset*  $I$ , bentuk semua subset yang tak kosong dari  $I$



- 2-itemset  $\{X_{10}, X_{13}\}$

Semua *subset* yang tak kosongnya adalah  $\{X_{10}\}, \{X_{13}\}, \{X_{10}, X_{13}\}$ .

- 2-itemset  $\{X_{10}, X_{31}\}$

Semua *subset* yang tak kosongnya adalah  $\{X_{10}\}, \{X_{31}\}, \{X_{10}, X_{31}\}$ .

- 2-itemset  $\{X_{28}, X_{29}\}$

Semua *subset* yang tak kosongnya adalah  $\{X_{28}\}, \{X_{29}\}, \{X_{28}, X_{29}\}$ .

- 2-itemset  $\{X_{30}, X_{31}\}$

Semua *subset* yang tak kosongnya adalah  $\{X_{30}\}, \{X_{31}\}, \{X_{30}, X_{31}\}$ .

2. Untuk tiap *subset* yang tak kosong  $s$  dari  $I$ , bentuk aturan yang berbentuk " $s \Rightarrow (I - s)$ " dengan nilai *support* dan *confidence*-nya, dengan syarat  $I - s \neq \emptyset$ .

- Aturan 1 :  $X_{10} \Rightarrow X_{13}$   $support(X_{10} \Rightarrow X_{13}) = 0.26$

$$confidence(X_{10} \Rightarrow X_{13}) = 0.65$$

- Aturan 2 :  $X_{13} \Rightarrow X_{10}$   $support(X_{13} \Rightarrow X_{10}) = 0.26$

$$confidence(X_{13} \Rightarrow X_{10}) = 0.76$$

- Aturan 3 :  $X_{10} \Rightarrow X_{31}$   $support(X_{10} \Rightarrow X_{31}) = 0.29$

$$confidence(X_{10} \Rightarrow X_{31}) = 0.72$$

- Aturan 4 :  $X_{31} \Rightarrow X_{10}$   $support(X_{31} \Rightarrow X_{10}) = 0.26$   
 $confidence(X_{31} \Rightarrow X_{10}) = 0.81$
- Aturan 5 :  $X_{28} \Rightarrow X_{29}$   $support(X_{28} \Rightarrow X_{29}) = 0.31$   
 $confidence(X_{28} \Rightarrow X_{29}) = 0.67$
- Aturan 6 :  $X_{29} \Rightarrow X_{28}$   $support(X_{29} \Rightarrow X_{28}) = 0.31$   
 $confidence(X_{29} \Rightarrow X_{28}) = 0.75$
- Aturan 7 :  $X_{30} \Rightarrow X_{31}$   $support(X_{30} \Rightarrow X_{31}) = 0.29$   
 $confidence(X_{30} \Rightarrow X_{31}) = 0.76$
- Aturan 8 :  $X_{31} \Rightarrow X_{30}$   $support(X_{31} \Rightarrow X_{30}) = 0.29$   
 $confidence(X_{31} \Rightarrow X_{30}) = 0.81$

#### 4.4.3 Pemilihan Aturan-Aturan yang Kuat

Dari hasil pencarian aturan-aturan asosiasi di atas, dapat dilihat dalam penelitian ini dihasilkan delapan asosiasi, suatu aturan asosiasi disebut kuat apabila memiliki *confidence* lebih besar atau sama dengan *confidence minimum* = 0.65. Dari keterangan di atas, semua aturan asosiasi yang didapat adalah aturan asosiasi yang kuat, karena semua aturan asosiasi tersebut memiliki *confidence* yang lebih besar atau sama dengan 0.65.

#### 4.4.4 Interpretasi Setiap Aturan Asosiasi yang Kuat

Setelah mendapatkan aturan-aturan asosiasi yang kuat, dilakukan interpretasi pada tiap aturan asosiasi yang kuat tersebut. Interpretasi ini berguna untuk memberikan informasi mengenai kombinasi *item* yang sering dibeli oleh pelanggan minimarket yang bersangkutan.

- Interpretasi Aturan 1 : “ *Jika seorang konsumen membeli mi instan, maka dengan kemungkinan 65 % ia akan membeli telur juga. Hal ini cukup signifikan karena mewakili 26 % dari database yang ada*”.
- Interpretasi Aturan 2 : “ *Jika seorang konsumen membeli telur, maka dengan kemungkinan 76 % ia akan membeli mi instan juga. Hal ini cukup signifikan karena mewakili 26 % dari database yang ada*”.
- Interpretasi Aturan 3 : “ *Jika seorang konsumen membeli mi instan, maka dengan kemungkinan 72 % ia akan membeli sikat gigi juga. Hal ini cukup signifikan karena mewakili 29 % dari database yang ada*”.
- Interpretasi Aturan 4 : “ *Jika seorang konsumen membeli sikat gigi, maka dengan kemungkinan 81 % ia akan membeli mie instan juga. Hal ini cukup signifikan karena mewakili 29 % dari database yang ada*”.
- Interpretasi Aturan 5 : “ *Jika seorang konsumen membeli sabun mandi, maka dengan kemungkinan 67 % ia akan membeli sabun cuci juga.*”

*Hal ini cukup signifikan karena mewakili 31 % dari database yang ada”.*

- Intepretasi Aturan 6 : “ *Jika seorang konsumen membeli sabun cuci, maka dengan kemungkinan 75 % ia akan membeli sabun mandi juga. Hal ini cukup signifikan karena mewakili 31 % dari database yang ada”.*
- Intepretasi Aturan 7 : “ *Jika seorang konsumen membeli pasta gigi, maka dengan kemungkinan 76 % ia akan membeli sikat gigi juga. Hal ini cukup signifikan karena mewakili 29 % dari database yang ada”.*
- Intepretasi Aturan 8 : “ *Jika seorang konsumen membeli sikat gigi, maka dengan kemungkinan 81 % ia akan membeli pasta gigi juga. Hal ini cukup signifikan karena mewakili 29 % dari database yang ada”.*

#### **4.5 HASIL PENELITIAN**

Dengan menetapkan nilai *minimum support* sebesar 0.25 dan *minimum confidence* sebesar 0.65, didapat delapan buah aturan asosiasi dengan melibatkan enam *item*. Kombinasi *item* yang sering dibeli oleh pelanggan adalah (mi instant, telur), (mie instant, sikat gigi), (sabun mandi, sabun cuci), (pasta gigi, sikat gigi).