

### BAB III METODELOGI PENELITIAN

#### 3.1 Pengumpulan Data

1. Wawancara

Wawancara dilakukan dengan karyawan pihak manajemen dan *electronic data processing* (EDP) pada PD. Buana yang beralamat pada Jalan S. Parman No. 81 Purwokerto dengan tujuan untuk menggali informasi sebanyak – banyaknya mengenai sistem yang sedang berjalan.

2. Pengumpulan Data

Data yang dimaksud adalah data sekunder yang merupakan data transaksi harian produk di PD. Buana khususnya kebutuhan sehari – hari. Data yang digunakan terdiri dari tanggal, nomer transaksi, jumlah barang, nama barang, kode barang. Data transaksi dengan jumlah 139.747 *record* yang dilakukan selama 3 bulan yaitu dari bulan Mei, Juni, dan Juli.

#### 3.2 Data Selection

Data yang diperoleh baik dari *database* maupun laporan transaksi perusahaan, memiliki isian – isian yang tidak sempurna seperti data yang hilang, data yang kosong, dan data yang tidak valid atau hanya sekedar salah ketik. Data yang tidak relevan akan dibuang karena keberadaanya bisa mengurangi mutu atau akurasi dari hasil data mining. Pada Tabel 3.1, menunjukkan terdapatnya isian atribut yang tidak valid.

Tabel 3.1 Data Selection

Tanggal	Nomer Transaksi	Jumlah Barang	Nama Barang	Kode Barang
1/7/2016	4237676788		IL SOAP B2 CLASSIC 100GR 36 X 100	5059
1/7/2016	3767689	1	CB H&B WASH 600+150XF S&S 12 X 750ML	46355

Tanggal	Nomer Transaksi	Jumlah Barang	Nama Barang	Kode Barang
1/7/2016	4237676789	1	CB HAIR&BODY WASH200MLMG 12 X 200 ML	41855
1/7/2016	4237676789	3	CB MB 200ML FRESH&NOURISH 12 X 200 ML	41851
1/7/2016	4237676790	1	CB CREAM 50GR C&P TUBE 50 X 48	Null
1/7/2016	4237676790	1	S17 SPRAY COLOGNE BRZNEW 48 X 75ML	40712
1/7/2016	4237676791	8		88607
1/7/2016	4237676792	1	DETOL ID HK 105 ENERGIZE 144 X 105GR	95104

### 3.3 Preprocessing Data

Data yang kosong atau null dihilangkan agar relevan atau sesuai dengan target atau output yang akan dihasilkan. Penyesuaian data harus menghasilkan data yang sesuai agar mudah dalam melakukan tahapan selanjutnya. Pada Tabel 3.2, menunjukkan atribut dengan isian yang sudah relevan.

Tabel 3.2 Preprocessing Data

Tanggal	Nomer Transaksi	Jumlah Barang	Nama Barang	Kode Barang
1/7/2016	4237676788	1	IL HW BOT PAMPERING PMP 12 X 250ML	41428
1/7/2016	4237676789	1	CB H&B WASH 600+150XF S&S 12 X 750ML	46355

Tanggal	Nomer Transaksi	Jumlah Barang	Nama Barang	Kode Barang
1/7/2016	4237676789	1	CB HAIR&BODY WASH200MLMG 12 X 200 ML	41855
1/7/2016	4237676789	3	CB MB 200ML FRESH&NOURISH 12 X 200 ML	41851
1/7/2016	4237676790	1	S17 SPRAY COLOGNE BRZNEW 48 X 75ML	40712
1/7/2016	4237676790	1	S17 SPRAY COLOGNE BRZNEW 48 X 75ML	40712
1/7/2016	4237676791	1	SWEETY PANTZ HK M 18S 12 X 18	88244
1/7/2016	4237676792	1	DETOL ID HK 105 ENERGIZE 144 X 105GR	95104

### 3.4 Transformed Data

Pada proses transformasi data merupakan data yang telah dipilih untuk melakukan proses pengolahan data. Pada Tabel 3.3, menunjukkan bahwa atribut yang digunakan adalah nomor transaksi, jumlah barang, dan kode barang. Pemilihan atribut nomor transaksi berdasarkan jumlah konsumen yang berbelanja, pemilihan atribut jumlah barang didasarkan pada banyaknya jumlah suatu produk yang dibeli konsumen dan pemilihan atribut kode barang didasarkan pada kode produk/barang.

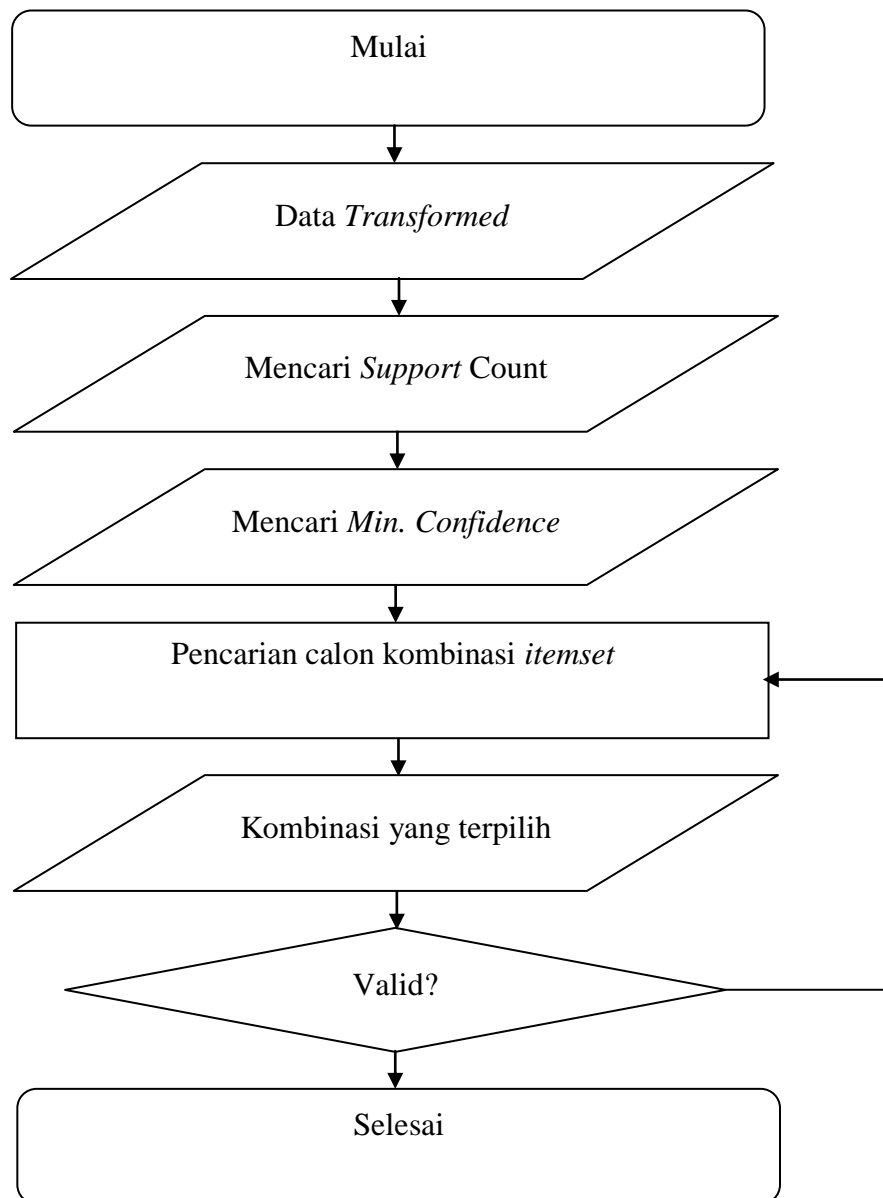
Tabel 3.3 *Transformed Data*

Nomer Transaksi	JumlahBarang	Kode Barang
4249277397	1	41428
4249277397	1	46355
4249277397	1	41855
4249280405	3	41851
4249280405	1	40712

Nomer Transaksi	JumlahBarang	Kode Barang
4249280405	1	40712
4249280405	1	88244
4249280406	1	95104

### 3.5 Algoritma A-Priori

Data yang sudah dipilih kemudian dianalisis menggunakan algoritma *A-Priori* untuk menemukan pola – pola dari *item* yang saling berhubungan. Pada Gambar 3.1, menjelaskan sebuah *flowchart* pada algoritma *A-Priori*.



Gambar 3.1 *Flowchart* Algoritma *A-Priori*

Proses algoritma *A-Priori* secara detail, berdasarkan Gambar 3.1 *Flowchart* Algoritma *A-Priori*.

1. *Data Transformed*

Menyeleksi dan membersihkan data – data yang akan dianalisis. Data *transformed* yang digunakan sebanyak 12.545 *record* transaksi. Atribut yang digunakan adalah nomor transaksi, jumlah barang, dan kode barang.

2. Mencari *Support Count*

Tentukan *support*, seleksi beberapa *item* data yang memenuhi minimal *support* sesuai dengan jumlah *item* barang dalam transaksi.

$$Support(A \cap B) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Total Transaksi}} \dots\dots\dots (3.1)$$

3. Mencari *Minimum Confidence*

Tentukan *confidence*, seleksi beberapa *item* data yang memenuhi minimal *confidence* sesuai dengan jumlah *item* barang dalam transaksi.

$$Confidence = P(B| A) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Jumlah Transaksi yang mengandung A}} \dots\dots\dots (3.2)$$

4. Pencarian calon kombinasi *itemset*

Mencari semua jenis *item* nama barang yang ada di *list* transaksi penjualan dan mencari jumlah setiap *item* yang ada pada setiap transaksi penjualan.

5. Kombinasi yang terpilih

*Item – item* yang telah terseleksi dibentuklah kombinasi 2 *item*, maka terbentuk beberapa *item* data dengan kombinasi 2 *item* yang berbeda, dengan *support* yang ditentukan maka terseleksi beberapa data 2 *item* (kombinasi 2 *item*). Kombinasikan *item – item* sampai batas maksimal *item* transaksi.

**3.6 Interpretation and Evaluation**

Pola yang telah diidentifikasi oleh sistem ke dalam bentuk yang lebih mudah dimengerti untuk membentuk pengambilan keputusan. Misalnya menunjukkan *item* yang saling berasosiasi melalui tabel atau grafik yang lebih mudah dimengerti. Berdasarkan hasil analisis pola keranjang belanja diharapkan dapat meningkatkan strategi – strategi untuk mempertahankan perusahaan serta meningkatkan penjualan.

Tahapan analisis ini menggunakan nilai *lift ratio*. *Lift ratio* digunakan untuk mengevaluasi kuat tidaknya sebuah aturan asosiasi. Untuk menghitung *lift ratio* digunakan rumus sebagai berikut:

$$Lift\ Ratio = \frac{Confidence}{Benchmark\ Confidence} \dots\dots\dots (3.3)$$

Untuk mendapatkan nilai *benchmark confidence* sendiri dapat dihitung menggunakan rumus sebagai berikut:

$$Benchmark\ Confidence = \frac{Nc}{N} \dots\dots\dots (3.4)$$

Keterangan:

Nc = jumlah transaksi dengan item yang menjadi *consequent*

N = jumlah transaksi basis data.