

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Gambaran Umum Objek Penelitian**

Objek penelitian yang digunakan dalam penelitian ini adalah Toko M Aji yang terletak di jalan SM Amin Pekanbaru, Riau. Toko ini merupakan usaha yang bergerak dibidang penjualan alat tulis kantor. Target pasar utama dari penjualan ATK pada toko ini adalah anak sekolah, pelajar mahasiswa-mahasiswi, dan kantoran. Selain itu, Toko M Aji juga merupakan *supplier* dari toko-toko ATK yang menjual barang satuan, hal ini dikarenakan sistem penjualan barang di toko M Aji dijual secara grosiran.

#### **3.2 Metode Penelitian**

Metode yang digunakan pada penelitian ini adalah metode kuantitatif. Metode kuantitatif dipilih karena melibatkan pemanfaatan analisis data numerik untuk menjelaskan suatu masalah ataupun fenomena melalui pengumpulan data yang berbentuk numerik. Metode ini digunakan untuk mendapatkan data pada masa lampau, mengenai transaksi penjualan barang alat tulis kantor di toko ATK M Aji.

##### **3.2.1 Metode Penyelesaian**

Pada penelitian ini akan diselesaikan dengan menerapkan *data mining* kemudian diproses dengan algoritma Apriori dan ECLAT dengan mencari asosiasi terhadap barang-barang yang dijual. Untuk data yang akan diteliti akan diproses dengan menggunakan *tools* RStudio.

##### **3.2.2 Alur Penelitian**

Alur penelitian yang akan dilakukan pada penelitian ini yaitu diambil berdasarkan perbandingan dari *process model data mining* yang sering

digunakan, yaitu KDD dan CRISP-DM. Berikut adalah perbandingan dari *process model* tersebut:

Tabel 3. 1 Perbandingan Proses Model KDD dan CRISP-DM

<i>Process Model</i>	<b>Alur proses</b>
KDD	<ol style="list-style-type: none"> <li>1. Alur proses lebih condong ke riset</li> <li>2. Belum diketahui siapa yang mau menggunakannya, karena berfokus pada <i>data mining</i> yang lebih mencari pengetahuan (<i>knowledge</i>).</li> </ol>
CRISP-DM	<ol style="list-style-type: none"> <li>1. Alur proses lebih condong ke industri</li> <li>2. Sudah diketahui siapa yang mau menggunakannya, karena sudah ada aturan yang ditetapkan dahulu setelah itu baru dicari datanya.</li> </ol>

Dari perbandingan pada tabel 3.1, dipilih *process model data mining* yang lebih cocok digunakan untuk penelitian ini adalah CRISP-DM, karena penelitian ini menggunakan data transaksi industri perdagangan jual beli ATK. Berikut adalah alur penelitian berdasarkan *process model* CRISP-DM:

a. *Business Understanding*

*Business understanding* adalah mengubah permasalahan bisnis menjadi permasalahan data analisis, yaitu membuat aturan asosiasi untuk pengadaan promosi dengan menggunakan algoritma Apriori dan algoritma ECLAT.

b. *Data Understanding*

Pengumpulan data sekunder dilakukan dengan cara mengambil data laporan transaksi penjualan pada toko ATK M Aji periode 2 bulan yang diambil dari tanggal 1 Desember 2021 sampai dengan tanggal 31 Januari 2022.

c. *Data Preparation*

Data yang diperoleh dari toko M Aji akan dilakukan dengan *data cleansing*. Data transaksi lalu dibersihkan dengan menghapus atribut-atribut yang tidak dibutuhkan dan digunakan pada penelitian yang dilakukan. Selanjutnya data yang sudah melalui tahap *cleansing* disimpan dalam bentuk format CSV untuk dianalisis menggunakan *tools* RStudio.

d. *Modeling*

Pembuatan model akan dilakukan dengan menggunakan algoritma Apriori dan ECLAT, untuk membuktikan berdasarkan perbandingan kinerja analisis

tiap algoritma. Pada algoritma apriori dan ECLAT batas *minimum support* (minsupp), *minimum confidence* (minconf), *minimum length* (minlen), *maximum length* (maxlen), dan *minimum lift rasio* (*lift*) dapat ditentukan sendiri. Penelitian dilakukan dengan percobaan terhadap keseluruhan data transaksi dengan nilai *minimum support* dan nilai *minimum confidence* yang berbeda, mulai dari nilai *minimum support* tertinggi dan nilai *minimum confidence* terendah. Nilai *minimum length* yang digunakan adalah 2 dan 3, *minimum support* yang digunakan adalah 10% sampai dengan 100%, dan untuk nilai *minimum confidence* juga dimulai dari 10% sampai dengan 100%. Tujuan dilakukannya pengujian dengan nilai *minimum support* dan *minimum confidence* yang berbeda adalah untuk membandingkan setiap hasil pengujian pada data transaksi penjualan di Toko ATK M Aji.

e. *Evaluation*

Tahap evaluasi yaitu untuk membandingkan *rules* yang diperoleh dengan menggunakan algoritma Apriori dan *rules* yang diperoleh dengan menggunakan algoritma ECLAT. Kualitas yang dibandingkan adalah berbentuk *minimum support*, *minimum confidence*, *lift ratio*, waktu, dan pendapat dari pemilik dan pegawai Toko M Aji mengenai hasil penelitian.

f. *Deployment*

Untuk tahap ini tidak termasuk, sebagaimana yang sudah dijelaskan pada batasan masalah.

### 3.3 Variabel Penelitian

#### 3.3.1 Variabel Dependen

Variabel Dependen yaitu variabel yang membuat terjadinya variabel independent dan merupakan variabel yang dijelaskan atau dipengaruhi oleh variabel independen. Variabel dependen yang terlibat yaitu:

Y1 = Aturan Asosiasi

### 3.3.2 Variabel Independen

Variabel Independen merupakan variabel bebas yang mempengaruhi atau yang menjadi sebab dari variabel dependen. Berikut adalah variabel independent yang ada:

X1 = Tanggal

X2 = Nama\_Pelanggan

X3 = Nama\_Item

### 3.4 Teknik Pengumpulan Data

Pada teknik pengumpulan data, data yang digunakan adalah data sekunder yang diambil langsung dari hasil wawancara dengan pemilik toko ATK M Aji untuk mengetahui bisnis toko tersebut. Kemudian wawancara dilakukan untuk mendapatkan *requirement* atau kebutuhan pada *management* toko.

### 3.5 Teknik Pengambilan *Sample*

Teknik pengambilan *sample* dilakukan dengan cara mengambil data laporan transaksi penjualan pada toko ATK M Aji pada periode 2 bulan yang diambil dari tanggal 1 Desember 2021 sampai dengan tanggal 31 Januari 2022. Data yang didapatkan cukup untuk diteliti karena memiliki 35.565 data transaksi.

### 3.6 Teknik Analisis Data

#### 3.6.1 Tools yang Digunakan

*Tools* yang digunakan untuk pengolahan data dan analisa data dilakukan pada penelitian ini diambil berdasarkan perbandingan *tools* yang sering digunakan untuk melakukan implementasi algoritma pada *data mining* yaitu *RapidMiner* dan RStudio. Berikut perbandingan dari *tools* tersebut:

Tabel 3. 2 Perbandingan *Tools Data Mining*

<b><i>Tools</i></b>	<b>Kelebihan</b>
<i>RapidMiner</i>	<ol style="list-style-type: none"> <li>1. Merupakan <i>tools</i> yang menggunakan sistem <i>drag and drop</i>.</li> <li>2. Dapat menampilkan visualisasi namun tidak dapat menggunakan <i>interactive visualization</i></li> <li>3. Dalam menganalisa hanya melakukan proses <i>input</i>.</li> </ol>
RStudio	<ol style="list-style-type: none"> <li>1. Merupakan <i>tools</i> yang memiliki <i>language support</i> dalam penggunaannya.</li> <li>2. Sangat baik dalam visualisasi data karena ada banyak <i>packages</i> yang mendukung R untuk membangun visualisasi menarik yang dapat berbentuk <i>interactive visualization</i>.</li> <li>3. Dalam menganalisa dapat memasukkan <i>code</i> program.</li> </ol>

Berdasarkan Tabel 3.2, *tools* yang akan dipakai pada penelitian ini adalah RStudio yang mana *tools* tersebut memang diperuntukkan sebagai *software* analisa data mining yang berfokus pada statistik, analisis data, dan dapat menampilkan visualisasi secara *interactive*.

