BAB I
PENDAHULUAN

I.1 Latar Belakang Masalah

Distro Maternal Disaster merupakan tempat penjualan pakaian yang bertempat di Bandung sejak 2003. Produk utama yang dijual di Distro Maternal Disaster adalah pakaian. Selain pakaian, produk pendukung lainnya seperti tas, dompet, jaket, celana, dan topi yang dipasarkan di beberapa outlet yang tersebar.

Setelah melakukan wawancara dengan pihak distro Maternal Disaster, didapatkan informasi bahwa pihak distro selalu membuat produk baru setiap bulannya dan produk tersebut dikirimkan ke outlet yang telah tersebar. Berdasarkan survey kuesioner terhadap 14 outlet Maternal didapatkan informasi bahwa 11 dari 14 outlet menyatakan produk yang dikirimkan kesetiap outlet tidak sesuai dengan kebutuhan outlet. Ketidaksesusaian tersebut diantaranya karena jenis produk dan ukuran yang tidak sesuai dari setiap outlet, sehingga menjadikan setiap outlet memiliki kebutuhan yang berbeda terhadap produk. Dilihat dari hasil survey yang telah dilakukan, dapat diambil kesimpulan bahwa produk yang dikirimkan tidak sesuai dengan karakteristik outlet yang berbeda-beda sehingga outlet mengalami kesulitan dalam menjual produk Maternal.

Untuk mendapatkan informasi dari setiap outlet, maka diperlukan data penjualan dari setiap outlet pada bulan Agustus 2016 untuk dianalisis. Metode yang dapat digunakan dalam menganalisis banyak adalah metode *data mining*, yaitu metode yang digunakan untuk melihat pola atau set data yang berukuran besar[1]. Terdapat beberapa metode yang dapat digunakan tergantung dari tujuan yang akan dicapai. Salah satu metode yang dapat digunakan adalah metode *clustering*, yaitu metode yang memisahkan data kedalam sejumlah kelompok (*cluster*) menurut kemiripan karakteristik dari data tersebut.

Berdasarkan masalah yang telah dipaparkan sebelumnya, solusi yang mungkin dilakukan adalah dengan membentuk segmentasi outlet sehingga karakteristik kebutuhan produk dari setiap outlet dapat terlihat. Karakteristik outlet tersebut dapat terlihat berdasarkan data penjualan yang telah dilakukan.

Maka dari itu, segmentasi akan dilakukan berdasarkan data penjualan. Segmentasi outlet dapat menjadi pengetahuan dengan menganalisis karakteristik atau perilaku berbeda yang membutuhkan produk berbeda dari setiap segmentasinya. Sehingga diharapkan dapat memaksimalkan proses pengiriman produk sesuai dengan segmentasi outlet yang telah dilakukan.

Maka dari itu Maternal Disaster perlu menggunakan metode *data mining* dan *clustering* dalam menentuka segmentasi outlet dari data penjualan produk yang ada. Oleh karena itu dari penjelasan yang sudah dipaparkan diatas maka perlu dibuat suatu aplikasi "**Penerapan Data Mining pada Penjualan Produk di Distro Maternal Disaster Menggunakan Metode Clustering**".

I.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka didapatkan suatu perumusan masalah yaitu bagaimana cara menerapkan data mining dalam membentuk segmentasi outlet di Distro Maternal Disaster berdasarkan data penjualan menggunakan metode *clustering*.

I.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk menerapkan data *mining* dalam membentuk segmentasi outlet di Distro Maternal Disaster berdasarkan data penjualan menggunakan metode *clustering.*

Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah untuk membantu pihak Distro Maternal Disaster unutk membentuk segmentasi outlet sehingga mempermudah manager *promotion* dalam melakukan strategi bisnis yang tepat untuk karakteristik setiap outlet.

I.4 Batasan Masalah

Pembahasan permasalahan diharapkan tidak menyimpang dari pokok permasalahan, sehingga diperlukan batasan masalah. Adapun batasan masalah dari penelitian ini adalah sebagai berikut :

1. Data yang akan digunakan berupa file dengan format *.xls atau .xlsx.
2. Data yang dianalisis meliputi data penjualan pada bulan Agustus 2016.

3. Metode yang akan digunakan adalah metode *clustering* dengan algoritma yang digunakan adalah *Agglomerative Hierarchical Clustering* (AHC) [2].
4. Aplikasi yang akan dibangun berbasis *website*.

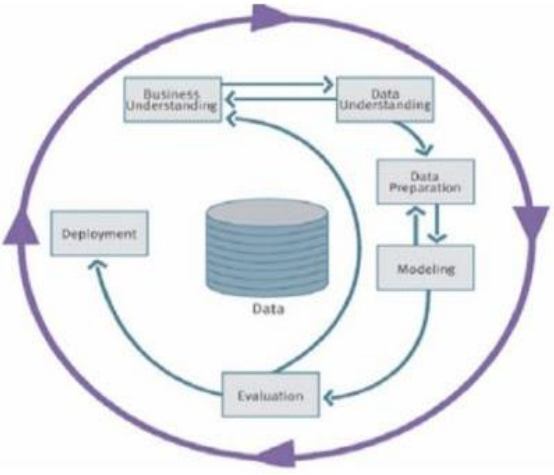
I.5 Metodologi Penelitian

Metode penelitian merupakan cara utama yang digunakan peneliti untuk mencapai tujuan dan menentukan jawaban atas masalah yang diajukan. Pada penelitian kali ini penulis akan menggunakan metode penelitian deskriptif.

Metode penelitian deskriptif adalah suatu metode yang digunakan untuk mencari unsur-unsur, ciri-ciri, sifat-sifat suatu fenomena. Metode ini dimulai dengan mengumpulkan data, menganalisis data, dan mengimplementasikannya. Metode deskriptif dalam pelaksanaannya dilakukan melalui teknik survey, studi kasus, studi komparatif, studi tentang waktu dan gerak, analisis tingkah laku, dan analisis dokumenter [3].

I.6 Metode Pembangunan Data *Mining*

Pada penelitian ini akan digunakan model proses yang mengadopsi dari model proses CRISP-DM (*Cross-Industry Standard Process for Data Mining*) untuk menyelesaikan masalah *data mining* [2]. Berikut ini merupakan tahapan penelitian yang akan dilakukan :



Gambar I. 1 Tahapan Penelitian

1. Pemahaman Bisnis

   Tahap pertama pada model proses CRISP-DM adalah tahap pemahaman bisnis (*business understanding*). Pada tahap ini, dilakukan pengumpulan informasi mengenai tujuan bisnis dan hasil yang diharapkan oleh manager distro Maternal Disaster. Kemudian dilakukan penyusunan rencana kegiatan yang akan dilakukan untuk mencapai tujuan yang telah ditentukan. Kegiatan pada tahap ini meliputi identifikasi tujuan bisnis, penilaian situasi, penentuan tujuan *data mining*, dan penyusunan rencana proyek.

   a. Identifikasi Tujuan Bisnis
      - Latar Belakang
      - Tujuan Bisnis
      - Kriteria Sukses Bisnis

   b. Penialian Situasi
      - Ketersediaan Sumber Daya
      - Ketersediaan Data
      - Ketersediaan Personel
      - Kebutuhan dan Asumsi

   c. Penentuan Tujuan Data *Mining*
      - Tujuan Data *Mining*
      - Kriteria Sukses Data *Mining*

2. Pemahaman Data

   Tahap kedua pada model proses CRISP-DM adalah tahap pemahaman data (*data understanding*). Pada tahap ini, dilakukan pengumpulan dan analisis terhadap data yang diperoleh dari distro Maternal Disaster. Kegiatan pada tahap ini meliputi pengumpulan data awal, pendeskripsian data, eksplorasi data, dan verifikasi kualitas data. Pada tahap ini akan dilakukan beberapa aktivitas, yaitu :

a. Pengumpulan Data Awal

   Pada tahap pengumpulan data awal akan dikumpulkan data-data yang akan digunakan dalam proses *clustering*.

b. Penjelasan Data

   Pada tahap ini data yang telah dikumpulkan akan dijelaskan apa saja isi dari data tersebut.

c. Eksplorasi Data

   Pada tahap ini data tersebut akan di eksplorasi untuk dapat membantu tujuan dari data *mining*.

d. Evaluasi Kualitas Data

   Pada tahap ini data tersebut akan di evaluasi kualitas datanya dengan pengecekan *missing value* dan *outlier*.

3. Persiapan Data

   Tahap ketiga pada model proses CRISP-DM adalah tahap persiapan data (data *preparation*). Pada tahap ini akan dilakukan pengolahan terhadap data yang telah diperoleh (data *preprocessing*). *Data preprocessing* perlu dilakukan untuk meningkatkan kualitas data agar dapat digunakan pada proses pemodelan, terdapat beberapa tugas dalam tahap ini, yaitu:

   a. Pemilihan Data

      Pada tahap ini akan dipilih atribut-atribut yang akan digunakan dalam proses data *mining*.

   b. Pembersihan Data

      Pada tahap ini data yang telah dipilih akan dibersihkan dari *missing value*.

   c. Penyusunan Data

      Pada tahap ini data yang telah dibersihkan dipersiapkan untuk digunakan pada pemodelan.

d. Integrasi Data

Pada tahap ini dilakukan penggabungan dua tabel atau lebih yang memiliki informasi berbeda tentang objek yang sama kedalam set data baru.

4. Pemodelan

Tahap pemodelan merupakan kegiatan utama dari proses penyelesaian masalah data *mining*. Data yang sudah disusun akan diproses menggunakan suatu metode sehingga dapat menghasilkan model. Kegiatan pada tahap pemodelan meliputi :

a. Teknik Pemodelan
b. Mekanisme Model
c. Penggunaan Teknik Pemodelan

5. Evaluasi

Evaluasi adalah tahap interpretasi terhadap hasil data *mining*. Evaluasi dilakukan secara mendalam denga tujuan agar hasil pada tahap *modelling* seuai dengan sasaran yang ingin dicapai dalam tahap *business understanding*.

a. Evaluasi Hasil
b. Peninjauan Proses
c. Menentukan Langkah Selanjutnya

6. Deployment

Deployment merupakan tahapan membuat laporan hasil kegiatan data *mining*. Laporan akhir mengenai pengetahuan yang didapat atau pengenalan pola pada data dalam proses data *mining* dan dipresentasikan dalam bentuk grafik atau deskriptif yang mudah dipahami.

I.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penyusunan laporan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini membahas mengenai Latar Belakang Masalah, Rumusan Masalah, Maksud dan Tujuan, Batasan Masalah, Metodologi Penelitian, dan Sistematika Penulisan yang digunakan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas tentang profil dari distro Maternal Disaster dan berbagai konsep dasar dan teori-teori yang berhubungan dengan judul penelitian tugas akhir.

BAB III ANALISIS DAN PERANCANGAN

Bab ini berisi tentang analisis dan perancangan sistem yang akan dibangun berdasarkan data transaksi yang diperoleh dari distro Maternal Disaster dengan menggunakan data *mining* dengan metode *clustering*.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan tentang implementasi dan analisis dari perancangan sistem yang dilakukan. Serta melakukan uji coba terhadap sistem yang telah di bangun dan hasil dari uji coba terhadap sistem tersebut

BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan dari hasil penelitian yang telah dilakukan dan saran-saran pengembangan lebih lanjut dari sistem yang dibangun.