

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang semakin pesat di era revolusi informasi saat ini telah berdampak besar pada semua bidang kehidupan, tidak terlepas pula pada bidang bisnis. Dengan adanya teknologi informasi menjadikan setiap *stakeholder* dapat mengakses berbagai data dan informasi yang dibutuhkan dengan mudah dan cepat. Dalam dunia bisnis, data dan informasi mempunyai peranan yang sangat penting. Data-data yang sangat detail mengenai produk, bahan baku, perilaku *end user* dan lain sebagainya sangat diperlukan dalam membuat strategi pemasaran. Disampaikan oleh Dan Zarella (2013), “*Marketing without data is like driving with your eyes closed*”, marketing yang dilakukan tanpa data maka ibarat menyetir dengan mata tertutup. Tanpa menggunakan data, strategi pemasaran yang dibuat tidak bisa mencapai tujuan bisnis dengan tepat.

Strategi pemasaran mengalami perubahan seiring dengan kemajuan teknologi. Semula strategi pemasaran berorientasi pada produk, kini strategi pemasaran juga harus mampu berorientasi pada pelanggan. Dengan strategi pemasaran yang berorientasi pelanggan, perusahaan dapat mengelola hubungan yang baik dengan pelanggan sehingga kepuasan pelanggan dapat tercapai (Shirole, 2021; Dasipah, 2017). Pencapaian kepuasan pelanggan yang tinggi akan berkaitan erat dengan perolehan laba, penghematan biaya, dan peningkatan pangsa pasar (Sager, 1994). Dalam penyusunan strategi pemasaran berorientasi pelanggan, langkah pertama yang perlu dilakukan adalah membuat segmentasi pelanggan. Segmentasi pelanggan dapat dimanfaatkan untuk mengidentifikasi karakteristik pelanggan dan mengelompokkan pelanggan berdasarkan perilaku atau kebiasaan yang sama (Brahmana, 2020; Shihab, 2019).

PT Aretha Nusantara Farm (ANF) adalah perusahaan yang bergerak di bidang peternakan, yang melakukan penjualan bibit ayam *broiler* atau yang lebih familier disebut dengan DOC (*Day Old Chick*) dan ayam *broiler* dewasa. PT ANF memiliki pelanggan yang tersebar di daerah Jawa Barat yaitu wilayah II

Purwakarta, Wilayah III Cirebon dan Wilayah IV Priangan, serta merambah juga ke daerah Jawa Tengah. PT ANF menyadari akan pentingnya hubungan antara pelanggan dengan keberhasilan bisnis. Pelanggan PT ANF yang datang dari berbagai daerah tentu memiliki karakteristik yang berbeda, hal ini menuntut PT ANF harus dapat membuat keputusan yang tepat dalam membuat strategi penjualan berdasarkan karakteristik pelanggan. Namun, dengan jumlah pelanggan yang tergolong banyak dan juga transaksi yang terus meningkat mengakibatkan PT ANF kesulitan dalam mengidentifikasi karakteristik pelanggan tersebut. Hal ini dapat mengakibatkan PT ANF kehilangan pelanggan potensial dan bisa merugikan perusahaan. Oleh karena itu, perlu dilakukan segmentasi pelanggan untuk mengidentifikasi karakteristik pelanggan dengan tepat.

Segmentasi pelanggan adalah mengelompokkan pelanggan yang memiliki kesamaan karakteristik sehingga dapat mengetahui perilaku (*behavior*) pelanggan (Shihab, 2019; Adiana, 2018). Perilaku pelanggan ini berulang sehingga membentuk pola, pola inilah yang harus dicermati oleh perusahaan sehingga bisa menjadi peluang untuk mengenal lebih jauh kepada para pelanggannya. Perusahaan dapat mengidentifikasi pola-pola pembelian pelanggan sehingga akan membantu dalam penerapan strategi pemasaran yang tepat untuk meningkatkan pendapatan perusahaan (Shirole, 2021; Muna, 2021).

Salah satu solusi dalam melakukan segmentasi pelanggan adalah dengan menggunakan teknik data mining (Shirole, 2021; Hung, dkk, 2019). Data mining adalah proses berulang pada data berskala besar untuk mengekstrak informasi dan pengetahuan yang akurat dan berguna untuk pengambilan keputusan (Vercellis, 2009). Dengan menggunakan metode-metode pada data mining, data yang diperoleh dari transaksi dapat dianalisa sehingga dapat menggali informasi-informasi tersembunyi dari data tersebut. Metode pada Data Mining yang digunakan untuk mengelompokkan data adalah metode *Clustering* (Aryuni, 2018; Sheshasaayee, 2018).

Metode *Clustering* dapat digunakan dalam menentukan segmentasi pelanggan dengan menganalisis kelompok data untuk mengetahui karakteristik dari kelompok pelanggan yang terbentuk (Wei, dkk, 2020). Terdapat dua tipe

metode *clustering* yaitu metode partisi dan metode hirarki. Metode partisi membagi data menjadi beberapa partisi yang saling terpisah, sedangkan metode hirarki memecah data menjadi hirarki klaster atau struktur pohon (Vercellis, 2009). Dalam penerapannya metode partisi memerlukan input jumlah klaster terlebih dahulu sedangkan pada metode hirarki tidak memerlukan input jumlah klaster. Metode partisi terdiri dari beberapa algoritma diantaranya adalah K-Means dan K-Medoids. Sedangkan untuk metode hirarki contohnya adalah algoritma Agglomerative dan Divisive (Vercellis, 2009).

Setiap metode *clustering* mempunyai kelebihan dan kekurangannya masing-masing. Hal tersebut menjadi fokus pada tujuan penelitian ini, yaitu untuk menentukan algoritma mana yang terbaik yang dapat digunakan pada data di PT ANF. Metode *clustering* partisi yang sangat populer dan banyak dipelajari karena sederhana dan mudah diterapkan adalah algoritma K-Means (Wei, dkk, 2020; Sheshasaayee; 2018; Jamal, 2019). Algoritma K-Means menggunakan nilai rata-rata (*mean*) sebagai pusat klaster (Mousavi, 2020; Muningsih, 2018). Metode *clustering* partisi lain yaitu Algoritma K-Medoids atau sering disebut juga dengan algoritma PAM (*Partitioning Around Medoid*) merupakan metode *clustering* yang cukup peka terhadap *outlier* (Mousavi, 2020) dan cukup efisien untuk data kecil (Muningsih, 2018). Metode K-Medoids menggunakan objek perwakilan (*medoid*) sebagai pusat *cluster* atau titik acuan pada setiap *cluster* (Aryuni, 2018). Sedangkan untuk metode *clustering* hirarki, algoritma yang paling populer digunakan adalah algoritma Agglomerative atau sering disebut *Agglomerative Hierarchical Clustering* (AHC). Metode AHC mempunyai kemampuan menggabungkan data dengan membuat hirarki, dimana yang memiliki kemiripan akan ditempatkan di hirarki yang berdekatan dan yang tidak memiliki kemiripan ditempatkan pada hirarki berjauhan (Parikh, 2020; Justitia, 2021).

Penerapan segmentasi pelanggan dengan metode *clustering* dapat menggunakan model RFM (*Recency, Frequency dan Monetary*). Model RFM ini diperkenalkan pertama kali oleh Arthur Huges pada tahun 1994. Model RFM ini merupakan proses penilaian berdasarkan perilaku pelanggan yang dilihat dari waktu transaksi terakhir pelanggan (*recency*), jumlah transaksi (*frequency*), dan

uang yang dikeluarkan (*monetary*) (Parikh, 2020; Monalisa, 2019). Dengan mempelajari rekaman interaksi dan transaksi pelanggan di masa lalu, perusahaan dapat melakukan penilaian terhadap tiga aspek dimensi tersebut. Model RFM dapat diaplikasikan untuk menganalisis perilaku pelanggan dan membangun strategi pemasaran (Shihab, 2019; Sutresno,dkk, 2018). Dalam penelitian ini dilakukan segmentasi pelanggan dari *dataset* transaksi berdasarkan karakteristik masing masing pelanggan pada perusahaan dan menganalisis segmentasi tersebut sehingga dapat memberikan rekomendasi untuk strategi pemasaran PT ANF.

Terdapat beberapa penelitian sebelumnya yang melakukan segmentasi pelanggan dengan menggunakan metode *clustering* dan model RFM. Penelitian yang dilakukan oleh Aryuni, dkk (2020) melakukan segmentasi nasabah pada bank XYZ dengan menggunakan algoritma K-means dan K-Medoids. Aryuni, dkk tidak menghilangkan data *outlier* pada proses persiapan data. Hasil penelitian menunjukkan bahwa metode K-Means mengungguli metode K-Medoids berdasarkan jarak intra *cluster* (AWC). Sementara berdasarkan indeks Davies-Bouldin, kinerja K-Means sedikit lebih baik daripada K-Medoids.

Penelitian lain yang dilakukan P.D. Hung (2019) menggunakan algoritma *hierarchical Agglomerative clustering* pada *dataset* kartu kredit untuk melakukan segmentasi pelanggan. P.D Hung menghilangkan data *outlier* pada proses persiapan data. Berdasarkan hasil yang diperoleh, peneliti dapat mempromosikan strategi pemasaran yang tepat yang lebih menguntungkan, namun kelemahan dari metode ini cukup lambat dan tergantung pada perangkat keras.

Penelitian yang dilakukan oleh Shihab, dkk (2019) melakukan segmentasi pelanggan dengan model RFM dan membandingkan algoritma K-Means dengan Agglomerative. Shihab tidak menghilangkan data *outlier* pada proses persiapan data. Hasil penelitian menunjukkan bahwa metode K-means lebih baik daripada Agglomerative berdasarkan dari waktu prosesnya sebesar 27,8% dan 97,8% serta berdasarkan jarak *intra cluster* dan jarak *inter cluster*.

Penelitian lain yang dilakukan oleh Shirole, dkk (2021) yang melakukan segmentasi pelanggan pada *e-commerce retail online* menggunakan model RFM dan algoritma K-Means. Shirole menghilangkan data *outlier* pada proses

persiapan data. Berdasarkan hasil yang diperoleh, segmentasi pelanggan yang terbentuk dapat membantu perusahaan untuk mengembangkan strategi pasar dan juga dapat digunakan sebagai media promosi kepada pelanggan setia mereka.

Berdasarkan penelitian-penelitian yang dilakukan sebelumnya dalam melakukan segmentasi pelanggan menggunakan metode *clustering*, terdapat perbedaan dalam melakukan persiapan data. Beberapa penelitian ada yang menghilangkan *outlier* pelanggan (P.D. Hung, 2019; Shirole, dkk, 2021), tetapi ada juga yang tidak menghilangkan *outlier* pelanggan (Aryuni, dkk, 2020; Shihab, 2019). Hal tersebut menjadi salah satu fokus pada tujuan penelitian ini, yaitu untuk menentukan bagaimana pengaruh *outlier* pada data dalam melakukan segmentasi pelanggan. Selain itu, penelitian ini juga berupaya menentukan algoritma mana yang terbaik dari metode *clustering* partisi dan hirarki yang dapat digunakan pada data di PT ANF. Sehingga diharapkan penelitian ini dapat menghasilkan segmentasi pelanggan yang tepat dan dapat digunakan untuk memberikan rekomendasi strategi pemasaran pada PT ANF. Oleh karena itu, penelitian ini diberi judul “Analisis Perbandingan Metode *Clustering* Partisi dan Hirarki dalam Menentukan Segmentasi Pelanggan Menggunakan Model RFM Pada PT Aretha Nusantara Farm”.

1.2 Identifikasi dan Rumusan Masalah

1.2.1 Identifikasi Masalah

Identifikasi masalah berdasarkan latar belakang yang telah disampaikan sebagai berikut:

1. Jumlah pelanggan yang tergolong banyak dan juga transaksi yang terus meningkat mengakibatkan PT ANF kesulitan dalam mengidentifikasi karakteristik pelanggan.
2. Terdapat beberapa metode *clustering*, yaitu metode *clustering* partisi dan metode *clustering* hirarki yang mempunyai kelebihan dan kekurangan masing-masing sehingga kesulitan dalam menentukan metode *clustering* mana yang tepat yang dapat digunakan sesuai dengan data di PT ANF.

3. Dalam melakukan segmentasi pelanggan dengan metode *clustering*, terdapat beberapa perbedaan pada penelitian saat melakukan persiapan data yaitu ada yang menghilangkan *outlier* dan ada yg tidak menghilangkan *outlier*.

1.2.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah:

1. Bagaimana perbandingan segmentasi pelanggan dengan menggunakan metode *clustering* partisi dan hirarki.
2. Bagaimana pengaruh *outlier* dalam melakukan segmentasi pelanggan dengan metode *clustering*.
3. Bagaimana rekomendasi strategi pemasaran PT ANF berdasarkan hasil segmentasi pelanggan dengan metode *clustering*.

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Untuk mengetahui perbandingan segmentasi pelanggan dengan menggunakan metode *clustering* partisi dan hirarki.
2. Untuk mengetahui pengaruh *outlier* dalam melakukan segmentasi pelanggan dengan metode *clustering*.
3. Untuk memberikan rekomendasi strategi pemasaran PT ANF berdasarkan hasil segmentasi pelanggan dengan metode *clustering*.

1.4 Manfaat Penelitian

1.4.1 Manfaat Pengembangan Ilmu

Manfaat penelitian ini dalam pengembangan ilmu yaitu sebagai berikut:

1. Dapat memberikan pengembangan keilmuan mengenai segmentasi pelanggan dengan menggunakan model RFM.
2. Dapat memberikan gambaran dalam memilih metode *clustering* partisi dan hirarki yang cocok dan tepat digunakan untuk segmentasi pelanggan di suatu perusahaan.

1.4.2 Manfaat Operasional

Manfaat penelitian ini untuk operasional perusahaan yaitu sebagai berikut:

1. Membantu PT ANF menentukan segmentasi pelanggan yang tepat sesuai karakteristik pelanggan
2. Memberikan rekomendasi strategi pemasaran pada PT ANF berdasarkan hasil segmentasi pelanggan.

1.5 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah data penjualan PT ANF selama 1 tahun, yaitu periode tahun 2021.
2. Metode *clustering* partisi yang digunakan adalah metode K-Means dan K-Medoids. Sedangkan metode *clustering* hirarki yang digunakan adalah Agglomerative.
3. Atribut yang digunakan dalam metode *clustering* menggunakan model RFM.

1.6 Sistematika Penulisan

Sistematika penulisan penelitian dapat dijelaskan sebagai berikut:

1. BAB I Pendahuluan

Bab ini berisi mengenai latar belakang penelitian, identifikasi dan rumusan masalah, tujuan penelitian, kegunaan penelitian, batasan masalah, dan sistematika penulisan.

2. BAB II Tinjauan Pustaka

Berisi tentang penjelasan dasar ilmu yang mendukung pembahasan penelitian ini meliputi: Metode *Clustering*, Algoritma K-Means, Algoritma K-Medoids, Algoritma Agglomerative, Model RFM (*Recency, Frequency, Monetary*). Selain itu berisi juga alur penelitian (*roadmap*) dari penelitian terdahulu dan kerangka penelitian.

3. BAB III Metodologi Penelitian

Metodologi penelitian menguraikan secara rinci pendekatan yang digunakan

dalam penelitian yaitu algoritma K-Means, algoritma K-Medoids dan algoritma Agglomerative dengan menggunakan atribut RFM (*Recency, Frequency, Monetary*) serta bantuan aplikasi Excel dan RStudio.

4. BAB IV Hasil Penelitian dan Pembahasan

Berisi tentang hasil penelitian dan pembahasan mengenai perbandingan metode *clustering* partisi dan hirarki dalam menentukan segmentasi pelanggan dengan menggunakan model RFM (*Recency, Frequency, Monetary*) di PT ANF.

5. BAB V Kesimpulan dan Saran

Berisi tentang kesimpulan dari hasil penelitian dan saran untuk mengembangkan penelitian yang telah dilakukan.