

BAB 2

TINJAUAN PUSTAKA

2.1 Landasan Teori

Landasan teori menjadi suatu dasar untuk menyusun tahapan pengerjaan penelitian. Memaparkan pokok pembahasan yang dilakukan dengan tujuan hasil yang sesuai.

2.1.1 PT Rohto Laboratories Indonesia

PT Rohto Laboratories Indonesia adalah anak perusahaan dari Rohto Pharmaceuticals Co., Ltd, perusahaan perawatan kesehatan konsumen asal Osaka, Jepang. Melahirkan produk-produk dalam bidang perawatan kesehatan dengan mutu tinggi dan dipercaya dunia seperti produk obat-obatan, perawatan kulit, perawatan bibir, perawatan rambut <https://www.rohto.co.id/>.

2.1.2 Business Intelligence

Business intelligence dapat digunakan untuk melakukan analisis data atau informasi bisnis yang disimpan kemudian diolah dengan jumlah data yang besar untuk memperoleh hasil dan melakukan pengambilan keputusan pada perusahaan dalam meningkatkan perkembangan bisnis (Ramadhan dkk. 2022). Penjelasan lebih lanjut mengenai *business intelligence* sebagai analisis, menyatukan dan mengakses data bisnis dalam jumlah besar untuk mendukung proses pengambilan keputusan dari data yang memberikan manfaat dengan informasi yang diperoleh dan mencapai serta mempertahankan keunggulan bisnis (Cahyati dkk. 2024).

Business intelligence bagi perusahaan akan memberikan peningkatan daya kerja seperti pemilihan strategi untuk bisnis, pelayanan dan mengatur anggaran yang baik. Dengan *business intelligence* juga membantu dalam mengolah informasi pada data yang tidak terstruktur menjadi lebih relevan dari kumpulan ide untuk kebutuhan bisnis. Hal ini mendukung dalam pengambilan keputusan dari informasi yang diperoleh (Nuryati dkk. 2023).

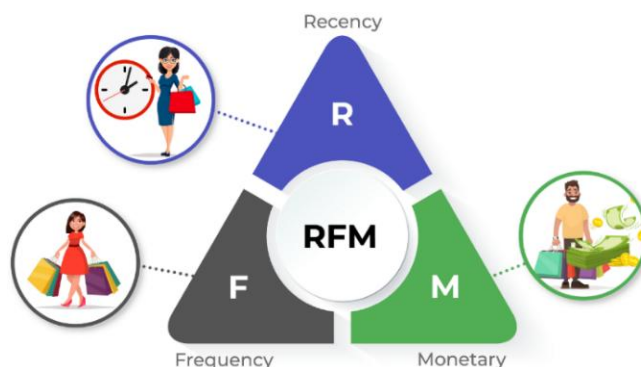
Penelitian ini menggunakan *business intelligence* karena cocok untuk mengolah data transaksi besar yang ada di PT Rohto Laboratories cabang Pekanbaru. Ini dapat membantu mengubah data menjadi informasi yang mudah

dipahami melalui visualisasi, sehingga pengguna dapat mengetahui perkembangan transaksi yang dilakukan oleh pelanggan pada tahun 2019 hingga 2024 bulan Agustus. Dengan begitu membantu dalam melakukan analisis dan membuat keputusan yang lebih tepat.

2.1.3 *Recency, Frequency, Monetary* (RFM)

Recency, frequency, monetary digunakan untuk memberi suatu keputusan pada pelanggan yang memiliki peluang untuk dianalisis pembelian terakhir yang dilakukan dari pelanggan disebut *recency*, untuk pelanggan yang sering melakukan transaksi disebut *frequency*, sedangkan *monetary* bagian pelanggan dengan total belanja yang tinggi pada periode tersebut (Wicaksana dkk., 2022). Penggunaan RFM dijadikan untuk analisis perilaku yang dilakukan oleh pelanggan terhadap transaksi. Ketika melakukan analisis *recency* menjadi variabel yang menghitung nilai pelanggan berdasarkan jarak waktu yaitu tanggal, bulan, tahun transaksi terakhir yang dilakukan pelanggan hingga saat ini, semakin kecil jarak waktu maka nilai *recency* menjadi besar. *Frequency* menjadi variabel yang menghitung nilai pelanggan dilihat dari jumlah transaksi yang dilakukan oleh pelanggan tersebut, semakin besar jumlah transaksi yang dilakukan maka nilai *frequency* yang diterima semakin besar. Sedangkan *monetary* menjadi variabel yang menghitung nilai pelanggan terhadap jumlah total pengeluaran yang dilakukan pelanggan dalam periode tersebut (Fadhillah dkk., 2025).

Model RFM menjadi metode efektif dalam melakukan segmentasi ketika dilihat dari data transaksi penjualan. Pelanggan akan dibagi menjadi beberapa bagian berdasarkan aktivitas pembelian dengan begitu dapat menyesuaikan berdasarkan karakteristik pelanggan (Al-Yasir dkk. 2024).



Gambar 2.1 Model RFM

Berdasarkan Gambar 2.1 nilai *Recency*, *Frequency*, *Monetary* (RFM) memiliki 4 karakteristik untuk segmentasi pelanggan seperti (Zahrani Putri dkk. 2023):

1) *Loyal Customer*

Kelompok pelanggan melakukan pembelian dengan jumlah transaksi dan jumlah pembayaran yang besar.

2) *Lost Customer*

Pelanggan yang sudah lama tidak melakukan transaksi, juga transaksi serta jumlah pengeluaran yang sedikit.

3) *Potential Customer*

Tergolong dalam pelanggan baru saja melakukan pembelian dengan jumlah dan pengeluaran yang masih sedikit.

4) *New Customer* tergolong dalam pelanggan baru saja melakukan pembelian dengan jumlah dan pengeluaran yang masih sedikit.

Berikut konsep perhitungan RFM yang digunakan (Andy Hermawan dkk., 2024):

$$R = SnapshotDate - TanggalTerakhirTransaksi$$

$$F = TotalJumlahTransaksi$$

$$M = \sum Total\ Nilai\ Transaksi$$

Setelah melakukan perhitungan nilai RFM yang diperoleh dari data transaksi yang dilakukan oleh pelanggan, penggunaan algoritma *K-Means* yang mengelompokkan pelanggan menjadi beberapa *cluster* yang telah ditentukan. Kombinasi dari model RFM dengan *K-Means* untuk membantu dalam analisis segmentasi pelanggan sudah banyak dikaitkan dengan berbagai penelitian. Hasil yang diperoleh pada kombinasi ini seperti membantu mengidentifikasi kelompok pelanggan, merancang strategi pemasaran yang lebih efektif, memperoleh hasil segmentasi pelanggan yang akurat serta bermanfaat dalam penentuan strategi yang tepat (Silamantha & Hadiono, 2024).

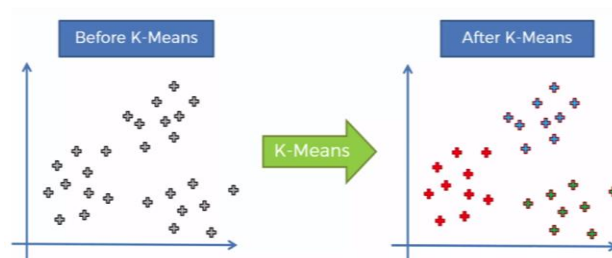
Penggunaan RFM pada penelitian ini membantu untuk mengetahui perilaku pelanggan berdasarkan transaksi yang dilakukan, seperti transaksi terakhir yang

dilakukan pelanggan, *frequency* pembelian serta total transaksi yang diberikan pelanggan untuk peningkatan terhadap PT Rohto Laboratories Pekanbaru. Tidak hanya sebagai dasar untuk segmentasi pelanggan, namun menjadi sumber informasi yang disimpan pada *data warehouse* kemudian divisualisasi dalam bentuk *dashboard business intelligence*. Hal ini akan membantu dalam proses analisis untuk memperoleh informasi pengambilan keputusan.

2.1.4 Segmentasi Pelanggan

Segmentasi menjadi bagian penting dalam proses penjualan dimana adanya hubungan dengan pelanggan menjadi lebih luas seperti memahami mengenai kebutuhan pelanggan. Segmentasi ini akan membuat pelanggan terbagi dalam beberapa *cluster* dengan kategori yang berbeda setiap pelanggannya. Segmentasi akan membangun strategi penjualan sebagai awal dalam pembuatan model bisnis (Harani dkk. 2020). Model yang banyak digunakan dan memiliki keakuratan yang tinggi yaitu *recency* dilihat dari hasil waktu terakhir pelanggan melakukan transaksi, *frequency* jumlah transaksi pelanggan peroleh, *monetary* banyaknya total transaksi yang dilakukan oleh pelanggan (Widiyanto & Witanti, 2021). Semakin kecil jangka waktu maka nilai *recency* semakin besar, semakin meningkat jumlah transaksi yang dilakukan maka nilai *frequency* semakin besar, semakin meningkat nominal yang dikeluarkan pelanggan maka *monetary* semakin besar (Fadhillah dkk., 2025). Adanya segmentasi pelanggan dalam *business intelligence* akan memperlihatkan karakteristik pelanggan berdasarkan data transaksi yang telah dilakukan. Dengan hasil segmentasi tersebut mempermudah dalam menentukan strategi penjualan yang sesuai berdasarkan pengelompokan pelanggan sehingga menjadi lebih efektif.

2.1.5 Algoritma K-Means Clustering



Gambar 2.2 K-Means Clustering

K-Means clustering sering digunakan dalam segmentasi pelanggan dan pengelompokan data seperti yang terdapat pada Gambar 2.2. Disebabkan kecepatan dan kemampuan perluasan pada *K-Means clustering* pada pengolahan data baik dan cocok untuk melakukan analisis data yang besar (Hendrastuty, 2024). *K-Means* memberi kemudahan dalam melakukan analisis dan mengkategorikan data untuk melihat transaksi penjualan jumlah yang besar dengan cepat dan tepat, serta pengelompokan data ke dalam beberapa tujuan tertentu (Afiasari dkk. 2023). *K-Means* memproses data yang telah disiapkan untuk membentuk beberapa kelompok berdasarkan kemiripan karakteristik. Hasil pengelompokan tersebut untuk melakukan analisis segmentasi pelanggan dan mendukung penyusunan strategi penjualan.

Pengelompokan algoritma *K-Means* dilihat dari kesamaan objek dengan kelompok lainnya. *K-Means* ini dapat digunakan dalam berbagai hal karena penerapan sederhana, mudah untuk diterapkan dan mampu untuk membuat pengelompokan data yang besar (Ahsina dkk. 2022). Dengan penggunaan algoritma *K-Means* perlu adanya metode *Elbow* membantu dalam menentukan jumlah *cluster* sehingga menghasilkan *cluster* dengan kualitas yang baik. Metode *Elbow* menunjukkan bentuk titik sudut siku dengan *cluster* sehingga semakin besar selisih nilai, maka menghasilkan jumlah *cluster* yang baik (Riani dkk. 2023). Proses dasar algoritma *K-Means* dapat dilihat di bawah ini (Dwi Agustiar dkk. 2025):

- 1) Jumlah *cluster* yang harus ditentukan terlebih dahulu berdasarkan kebutuhan yang diperlukan.
- 2) Ukuran jarak menjadi tahapan penting untuk menempatkan observasi berdasarkan nilai *centroid*.
- 3) Hitung *Euclidean Distance* antar titik data ke setiap pusat *cluster* dengan persamaan berikut:

$$d(p, q) = \sqrt{((x_2 - x_1)^2 + \dots + (z_n - z_{n-1})^2)}$$

untuk $(x_2, \dots, z_n)$ sebagai koordinat pusat *cluster* dan $(x_1, \dots, z_{n-1})$ sebagai koordinat titik data.

- 4) Hitung *Euclidean Distance* antar titik data ke setiap pusat *cluster* dengan persamaan berikut:

$$v = \frac{\sum_{i=1}^n X_i}{n}$$

keterangan

v : Pusat *cluster*

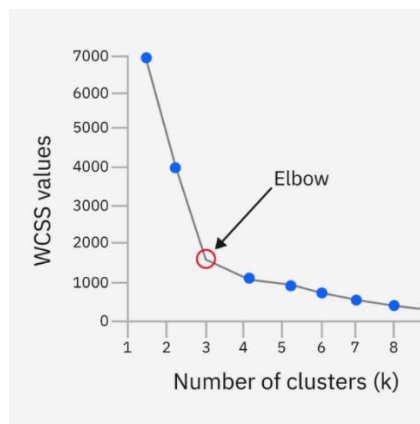
X_i : Titik data ke- i

n : Banyaknya titik data dalam satu *cluster*

- 5) Ulangi langkah ke-3 dan ke-4 hingga tidak ada perubahan pada pengelompokan titik data atau sampai mencapai kriteria berhenti yang diterapkan.

K-Means digunakan sebagai alat analisis dalam mengelompokkan pelanggan berdasarkan karakteristik dari perhitungan RFM. Penggunaan algoritma ini menjadi salah satu bagian analisis dalam *business intelligence*. Hasil yang diperoleh dari pengelompokan pelanggan digabungkan ke dalam *data warehouse* dan ditampilkan pada *dashboard* untuk digunakan sebagai dasar pembuatan strategi penjualan.

2.1.6 Elbow Method



Gambar 2.3 Elbow Method

Elbow Method menjadi metode yang digunakan dalam algoritma *K-Means Clustering* sebagai pengujian kualitas segmentasi *K-Means* untuk menentukan jumlah *cluster* yang terbaik. Metode *elbow* akan menghitung total *Within-Cluster Sum of Squared* (WCSS) untuk setiap nilai k , yang menjadi ukuran seberapa dekat data setiap kelompoknya dengan *centroid cluster* (Ryandra dkk. 2025). Berikut formula yang dapat digunakan untuk menghitung nilai *Within-Cluster Sum of*

Squared (WCSS) dalam menentukan jumlah *cluster*, sebagai berikut (Sigit dkk., 2025):

$$WCSS = \sum_{i=1}^K \sum_{x \in C_i} \|x - u_i\|^2$$

keterangan

- k : Jumlah Cluster
- x : Data dalam Cluster C_i
- u_i : Centroid dari Cluster C_i
- $\|x - u_i\|^2$: Jarak Euclidean kuadrat antara data dan centroid

2.1.7 Data Warehouse

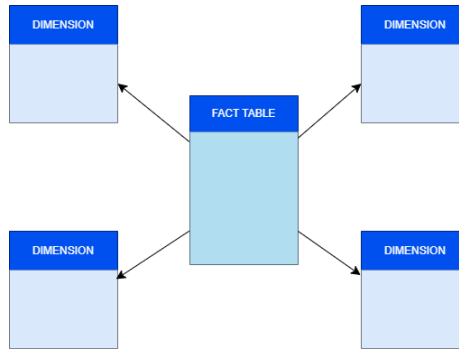
Data warehouse digunakan sebagai solusi untuk meringkas dan menyimpan data yang digunakan untuk analisis serta tujuan menyediakan data dengan informasi akurat dan cepat dalam mendukung pengambilan keputusan yang lebih baik (Firliana dkk. 2024). *Data warehouse* dapat mencakup informasi dengan berbagai waktu, penyimpanan data yang disesuaikan pada subjek, terintegrasi, dan tidak mudah berubah ketika proses pengumpulan data dalam pengambilan keputusan (Nur Fauziah dkk. 2024).

1) Dimensi dan Tabel Fakta

Dimensi adalah struktur berisi informasi dimensi yang akan digunakan dalam *data warehouse* berdasarkan kebutuhan dalam analisis data. Dengan penggunaan dimensi mempermudah dalam tahapan analisis untuk mengetahui strategi penjualan sesuai tujuan penelitian ini (Eka Pratama & Bernard, 2022). Tabel fakta merupakan pusat tabel yang menampilkan data dalam bentuk angka maupun tambahan untuk analisis. Tabel fakta diperoleh dari hasil menghubungkan tabel dimensi (Hilmansyah & Maesaroh, 2022).

2) Star Schema

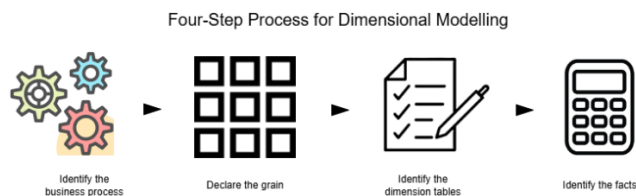
Star schema yang terlihat pada Gambar 2.4 adalah model data yang memiliki tabel fakta ditengah dengan tabel dimensi disekelilingnya. *Star schema* ditandai dengan satu atau lebih tabel fakta yang berisi informasi pada *data warehouse* dengan beberapa tabel dimensi yang lebih kecil serta masing-masingnya memiliki informasi atribut tertentu pada tabel fakta.



Gambar 2.4 Star Schema

Data warehouse pada penelitian ini menjadi bagian utama dalam arsitektur *business intelligence* sebagai tempat menyimpan data dalam jumlah besar dengan hasil penggabungan data yang berasal dari berbagai sumber. Tujuan penggunaan *data warehouse* untuk membantu dalam proses analisis dengan cepat yang menyediakan data yang siap digunakan untuk membuat keputusan. Dengan begitu dapat melihat informasi berdasarkan dimensi dan tabel fakta.

2.1.8 Metode *Four-Step Kimball*



Gambar 2.5 Four-Step Kimball

Pada metode *data warehouse* terdapat Gambar 2.5 *Four-Step Kimball* yang membantu dalam melakukan analisis data, dengan begitu memiliki 4 tahapan diantaranya (Lestari dkk., 2024):

1) Proses Bisnis

Pada langkah ini penetapan subjek utama. Subjek utama berpedoman pada proses bisnis yang dapat memberikan respon atau seluruh pertanyaan bisnis yang penting dan memiliki ciri khas.

2) *Grain*

Memutuskan sumber atau *grain* harus memilih gambaran atau pemaparan yang akan diutus oleh sebuah tabel fakta. Setelah *grain* dari tabel fakta ditentukan, tahapan berikutnya menetapkan tabel dimensi.

3) Identifikasi Dimensi

Langkah ini mengidentifikasi dan menyelaraskan dimensi dari tahapan yang ditentukan sebelumnya. Dimensi ini merupakan representasi dari bukti dalam tabel fakta yang akan digunakan.

4) Pemilihan Fakta

Pada tabel fakta terdapat data yang dapat dihitung atau *grain* yang dapat digambarkan melalui tabel atau grafik. Pada tahap ini, penetapan ukuran atau *measure* dibutuhkan dalam tabel fakta. Tabel fakta melingkupi *foreign key* dari tabel dimensi yang saling berkaitan.

Penggunaan metode *Four-Step Kimball* pada penelitian ini akan membantu dalam pembentukan struktur *star schema* yang mudah untuk dipahami dan sederhana. Hal ini memberi kemudahan untuk membuat tabel fakta serta tabel dimensi. Adanya *Four-Step Kimball* untuk membangun *data warehouse* sebagai sumber informasi dalam pembentukan visualisasi pada *dashboard business intelligence*.

2.1.9 MySQL

MySQL sebenarnya turunan dari konsep utama pada *database* yaitu SQL (*Structur Query Language*). SQL merupakan konsep untuk mengolah *database*, termasuk mengakses dengan melakukan seleksi dan pemasukan data yang memungkinkan dalam beroperasi secara otomatis. *Database server* ini dapat dikatakan unggul terbukti *query* dilakukan oleh *single user*, kecepatan *query* MySQL lebih cepat dari PostgreSQL dan Interbase (Zulfa & Wanda, 2023).

2.1.10 Laravel



Gambar 2.6 Laravel

Laravel merupakan *framework* yang mendukung untuk membangun *website* berbasis PHP yang memiliki keunggulan dalam kemudahan penggunaan, keamanan serta efisiensi. Laravel juga memiliki struktur kode yang tersusun dengan baik dan

juga terdapat fitur bawaan untuk mempermudah dalam pengembangan *web* seperti *routing*, *middleware* serta *sistem autentikasi*. Tidak hanya itu laravel juga memisahkan logika dari tampilan dengan begitu memberi kemudahan untuk dalam olah kode seperti penerapan *MVC* atau model, *view* dan *controller* (Sinlae dkk., 2024). Pada penelitian ini laravel membantu dalam pengembangan *dashboard business intelligence* berbasis *web* yang dihubungkan dengan MySQL sebagai tempat penyimpanan data, sehingga memperoleh hasil analisis yang dapat diakses dengan mudah oleh *user*.

2.1.11 Metode Pengujian

Black box testing menjadi salah satu metode untuk melakukan pengujian perangkat lunak di setiap bagian yang ada pada sistem. Pengujian dilakukan untuk memastikan bahwa semua fungsi dapat digunakan dengan baik tanpa menguji desain dan kode program dengan begitu mengetahui apakah fungsi, masukkan dan keluaran dari perangkat lunak sesuai dengan spesifikasi atau alur yang dibutuhkan (Yulianti & Prastowo, 2021).

User Acceptance Testing dilakukan oleh pengguna setelah *black box testing* yang berfokus pada fungsional sistem. Pengujian ini dilakukan untuk mengkonfirmasi hasil pengembangan perangkat lunak sudah sesuai atau belum berdasarkan kebutuhan pengguna, yang berfokus pada *input* dilakukan pengguna dan *output* yang dihasilkan sistem. Sehingga ketidaksesuaian yang mungkin dihasilkan dapat dievaluasi dan melakukan perbaikan (Dewi dkk. 2021).

2.2 Penelitian Terdahulu

Penelitian yang dilakukan oleh Damayanti & Putri (2024) bertujuan untuk melakukan analisis lebih lanjut dari pola penjualan beras dengan melihat segmen pasar berdasarkan karakteristik pembelian. Penelitian ini menggunakan *K-Means clustering* untuk mengetahui kelompok konsumen yang dimiliki dari pola pembelian beras dengan begitu akan mendapatkan strategi penjualan yang sesuai. Penelitian ini memiliki hasil akhir yang diimplementasikan ke dalam bentuk *web* dengan memperlihatkan data penjualan, halaman pengelompokan data.

Penelitian kedua dilakukan oleh Fajar dkk. (2024) bertujuan melakukan pengelompokan dari data penjualan obat berdasarkan perilaku konsumen untuk

memberikan peningkatan bisnis apotek An-Naafi. Penelitian ini menggunakan *K-Means* untuk menganalisis data penjualan obat dan menghasilkan rekomendasi untuk peningkatan strategi penjualan dalam bentuk *cluster*.

Penelitian ketiga dilakukan oleh Haston dkk. (2024) bertujuan untuk melakukan segmentasi pelanggan. Penelitian ini menggunakan *K-means clustering* untuk mengetahui karakter dan menganalisis pelanggan yang memiliki prospek tinggi. Dari penelitian akan memiliki hasil *cluster* yang membantu dalam menyesuaikan keinginan dari *customer* untuk manajemen produk baru yaitu *less sugar* dan menggunakan media *online* dalam pemasaran.

Penelitian keempat dilakukan oleh (Karami & Widharto, 2023) bertujuan untuk membuat sebuah sistem yang mempermudah dalam pemantauan kontribusi dari pelanggan. Adanya *business intelligence* dapat mengubah data menjadi informasi yang berguna dengan *K-Means clustering* yang mengelompokkan pelanggan berdasarkan kriteria tertentu, melakukan segmentasi pelanggan. Dari penelitian akan memiliki hasil efektif dalam membantu perusahaan untuk memahami kontribusi dari pelanggan.

Tabel 2.1 Penelitian Terdahulu

Kategori	Peneliti 1	Peneliti 2	Peneliti 3	Penelitian 4
Tujuan	Menganalisis pola penjualan beras dengan melihat segmen pasar berdasarkan karakteristik pembelian.	Pengelompokan dari data penjualan obat berdasarkan perilaku konsumen untuk memberikan peningkatan bisnis apotek An-Naafi.	Melakukan segmentasi dan karakteristik pelanggan.	Mempermudah pemantauan kontribusi menggunakan sistem
Algoritma	<i>K-Means Clustering</i>	<i>K-Means Clustering</i>	<i>K-Means Clustering (Unsupervised Learning)</i>	<i>K-Means Clustering</i>
Data	Penjualan beras	Penjualan obat 1 tahun terakhir sebanyak 1486 data	1000 orang yang berbelanja minuman ringan di India	Data sales
Hasil	Sistem berbasis web yang menampilkan <i>cluster</i>	Menganalisis data penjualan obat dan menghasilkan	<i>Cluster</i> yang membantu dalam menyesuaikan	Hasil efektif dalam membantu perusahaan

Kategori	Peneliti 1	Peneliti 2	Peneliti 3	Penelitian 4
	pelanggan yang tidak laris, laris, dan sangat laris maka, memperoleh 3 <i>cluster</i> penjualan.	rekomendasi untuk peningkatan strategi penjualan dalam bentuk <i>cluster</i> dengan membentuk 2 <i>cluster</i> .	keinginan dari <i>customer</i> untuk manajemen produk baru serta memperoleh 7 <i>cluster</i> .	untuk memahami kontribusi dari pelanggan

Penelitian terdahulu yang terdapat pada Tabel 2.1. membahas penggunaan algoritma *K-Means* sebagai segmentasi pelanggan dari segi pola penjualan, perilaku dan karakter pelanggan. Segmentasi yang dilakukan untuk pendukung dalam mengambil keputusan strategi penjualan dalam peningkatan bisnis. Sebagian besar penelitian sebelumnya berfokus pada penggunaan algoritma *K-Means* untuk mengetahui pengelompokan data meskipun adanya implementasi sistem berbasis *web*, tetapi tidak secara spesifik berfokus pada pengembangan sistem *business intelligence* berdasarkan segmen pelanggan dengan penggunaan RFM, *Four-Step Kimball* seperti yang dilakukan dalam penelitian ini sehingga hasil segmentasi pelanggan dapat dimanfaatkan untuk dasar pengambilan keputusan strategi penjualan. Meskipun terdapat penelitian yang menggunakan algoritma *K-Means* untuk segmentasi, pengembangannya hanya sebatas pada kelompok data tanpa adanya visualisasi dari hasil segmentasi dalam bentuk *dashboard*.