

BAB 4

PEMBAHASAN

4.1 *Data Mining Architecture*

Dalam proses *data mining*, *data set* yang digunakan diambil dari *data warehouse*. Namun dikarenakan *data warehouse* mencakup banyak proses bisnis keseluruhan dari perusahaan. Maka *data set* yang digunakan dalam penelitian ini diambil dari data transaksi penjualan pada *marketplace* Shopee dengan periode 3 bulan (1 oktober 2022 - 31 desember 2022). *Data set* kemudian dibandingkan polanya dengan *data training* dan *data set* akan di *testing* dengan pola yang sama.



Berikut adalah gambar arsitektur *data mining* dari penelitian ini :

Gambar 4.1 : Arsitektur *Data Mining* Penelitian

4.2 *Business Understanding*

Pada tahap awal ini, dilakukan pemahaman terhadap alur dari data transaksi terlebih dahulu serta penjelasannya terhadap *data mining* agar mencapai tujuan yang diinginkan. Pemahaman dalam penelitian ini mengacu pada provinsi-provinsi yang melakukan transaksi di akun Shopee Chibiland. Penjualan tersebut terbagi ke beberapa wilayah di Indonesia, namun jumlah penjualan disetiap wilayahnya berbeda. Maka dibutuhkannya analisis untuk mengetahui wilayah mana saja yang jumlah transaksinya lebih rendah dibanding provinsi lainnya. Sehingga perusahaan dapat mempertimbangkan membuat strategi penjualan khusus agar penjualan lebih merata disetiap wilayah di Indonesia.

4.3 *Data Understanding*

Data yang digunakan dalam penelitian ini yaitu data transaksi penjualan akun Shopee Chibiland pada periode Oktober – Desember 2022. Data ini ditarik dari akun Shopee Seller Chibiland yang dapat diakses pada <https://shopee.co.id/seller/login>. Dari data tersebut berisikan 413 records yaitu berisikan data dan status pesanan, data barang penjualan, data pembayaran dan data pemesan. Berikut data yang digunakan pada penelitian ini yang ditunjukkan pada gambar dibawah ini :

No. Pesanan	Status Pesanan	Alat Pembayaran	Status Pembatalan / Pengembalian	No. Resi	Opil Pengiriman	Alamat pengiriman	Pesanan Harus Dikirimkan Dobelem (Menghindari keterlambatan)	Waktu Pengiriman Datar	Waktu Pesanan Dibuat	Waktu Pembayaran Dibuat	SKU Instuk	Nama Produk	Nomor Referensi SKU	Nama Variasi	Harga Awal	Harga Setelah Diskon	Jumlah
20185H 6JPF9Q	Selesai			80251120 8642330	JNE Regular	Alamat Candi	18/10/2022 16:59	18/10/2022 16:38	10/5/2022 16:58	10/5/2022 16:59	3678203 75821	Cibiland Diany Frizon Set Pouch & Hand Puppel Etna	3578203 70021		329.000	128.688	1
20185H 6JPF9Q	Selesai			80251120 8642330	JNE Regular	Alamat Candi	18/10/2022 16:59	18/10/2022 16:38	10/5/2022 16:58	10/5/2022 16:59	3678203 75821	Cibiland Diany Frizon Set Pouch & Hand Puppel Anna	3578203 70021		329.000	128.688	1

Gambar 4.2 : Data penelitian dalam format Excel

4.4 Data Preparation

Tahap ini merupakan tahap dimana mulai melakukan persiapan data dengan menyesuaikan data mana saja yang dibutuhkan dalam membuat penelitian ini.

1) Data Selection

Dari data yang telah diambil didapatkan 413 transaksi penjualan selama periode Oktober – Desember 2022. Kemudian data ini diseleksi kembali dan hanya memilih No. Pesanan, Provinsi, Kota/Kabupaten dan Status Pesanan. Data hasil dapat dilihat pada gambar dibawah ini:

No. Pesanan	Provinsi	Kota/Kabupaten	Status Pesanan
201005H5JP6P8Q	JAWA TIMUR	KAB. SIDGARJO	Selesai
201005H5JP6P8Q	JAWA TIMUR	KAB. SIDGARJO	Selesai
201009UUMV53GM	JAWA BARAT	KOTA DEPOK	Selesai
201010V1CCJSJR	JAWA BARAT	KOTA SUKABUMI	Selesai
201010V1CCJSJR	JAWA BARAT	KOTA SUKABUMI	Selesai
201010VUNHA731	BANTEN	KOTA TANGSERANG	Selesai
20101005J0RYK9	JAWA TIMUR	KOTA SURABAYA	Selesai
2010100PCS17TX	KALIMANTAN BARAT	KOTA PONTIANAK	Selesai
2010100PD1RXBW	JAWA TIMUR	KOTA SURABAYA	Selesai
2010100PD1RXBW	JAWA TIMUR	KOTA SURABAYA	Selesai
2010100PD1RXBW	JAWA TIMUR	KOTA SURABAYA	Selesai

Gambar 4.3 : Data penelitian yang telah diseleksi

Setelah data diseleksi, data perlu dirapikan terlebih dahulu dengan menghitung jumlah status pesanan “Selesai” atau “Batal” disetiap provinsinya dengan melakukan *sort* terlebih dahulu agar data provinsi berurutan, kemudian menggunakan rumus *excel* yaitu COUNTIF untuk menghitung jumlah status pesanan “Selesai” atau “Batal” disetiap provinsi. Untuk lebih jelas dapat menggunakan Pivot Table.

Dan hasilnya akan seperti gambar dibawah ini:

Provinsi	Jumlah dari Transaksi Selesai	Jumlah dari Transaksi Batal
BALI	2	1
BANTEN	44	0
DI YOGYAKARTA	8	0
DI JAKARTA	107	18
JAWA BARAT	115	28
JAWA TENGAH	12	1
JAWA TIMUR	44	0
KALIMANTAN BARAT	2	2
KALIMANTAN TIMUR	2	2
KELULIAHMAN	2	0
LAMPUNG	2	0
NUA	2	0
SULAWESI UTARA	2	1
SUMATERA SELATAN	2	1
SUMATERA UTARA	2	1
SURABAYA	0	0
Total Keseluruhan	355	69

Gambar 4.4 : Data Pivot Transaksi

Hingga hasil akhir ditampilkan pada Tabel 4.1 Data Pivot Transaksi :

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
1	BALI	KAB. GIANYAR	1	1
2	BALI	KOTA DENPASAR	2	2
3	BANTEN	KAB. TANGERANG	9	9
4	BANTEN	KOTA CILEGON	1	1
5	BANTEN	KOTA SERANG	14	13
6	BANTEN	KOTA TANGERANG	11	11
7	BANTEN	KOTA TANGERANG SELATAN	9	7
8	DI YOGYAKARTA	KAB. BANTUL	1	1
9	DI YOGYAKARTA	KAB. KULON PROGO	5	3
10	DI YOGYAKARTA	KAB. SLEMAN	3	3
11	DKI JAKARTA	KOTA JAKARTA BARAT	46	38
12	DKI JAKARTA	KOTA JAKARTA PUSAT	17	16
13	DKI JAKARTA	KOTA JAKARTA SELATAN	19	15
14	DKI JAKARTA	KOTA JAKARTA TIMUR	18	7
15	DKI JAKARTA	KOTA JAKARTA UTARA	23	13
16	JAWA BARAT	KAB. BANDUNG	7	6
17	JAWA BARAT	KAB. BANDUNG BARAT	2	2
18	JAWA BARAT	KAB. BEKASI	12	11
19	JAWA BARAT	KAB. BOGOR	22	17

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
20	JAWA BARAT	KAB. CIANJUR	1	1
21	JAWA BARAT	KAB. INDRAMAYU	1	1
22	JAWA BARAT	KAB. KARAWANG	1	1
23	JAWA BARAT	KAB. KUNINGAN	1	1
24	JAWA BARAT	KAB. SUBANG	2	2
25	JAWA BARAT	KAB. SUKABUMI	1	1
26	JAWA BARAT	KAB. SUMEDANG	1	1
27	JAWA BARAT	KAB. TASIKMALAYA	1	1
28	JAWA BARAT	KOTA BANDUNG	11	10
29	JAWA BARAT	KOTA BEKASI	24	22
30	JAWA BARAT	KOTA BOGOR	6	5
31	JAWA BARAT	KOTA CIMAHI	2	2
32	JAWA BARAT	KOTA CIREBON	2	2
33	JAWA BARAT	KOTA DEPOK	37	32
34	JAWA BARAT	KOTA SUKABUMI	2	2
35	JAWA BARAT	KOTA TASIKMALAYA	3	3
36	JAWA TENGAH	KAB. BOYOLALI	1	1
37	JAWA TENGAH	KAB. BREBES	1	1
38	JAWA TENGAH	KAB. JEPARA	1	1
39	JAWA TENGAH	KAB. KARANGANYAR	1	1

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
40	JAWA TENGAH	KAB. KLATEN	2	2
41	JAWA TENGAH	KAB. PEMALANG	6	6
42	JAWA TENGAH	KAB. SRAGEN	1	1
43	JAWA TENGAH	KOTA SEMARANG	6	6
44	JAWA TIMUR	KAB. GRESIK	1	1
45	JAWA TIMUR	KAB. MALANG	1	1
46	JAWA TIMUR	KAB. SIDOARJO	3	2
47	JAWA TIMUR	KOTA KEDIRI	1	1
48	JAWA TIMUR	KOTA MADIUN	2	2
49	JAWA TIMUR	KOTA MALANG	8	8
50	JAWA TIMUR	KOTA SURABAYA	31	27
51	KALIMANTAN BARAT	KOTA PONTIANAK	3	3
52	KALIMANTAN TIMUR	KAB. KUTAI TIMUR	4	4
53	KEPULAUAN RIAU	KOTA BATAM	8	8
54	LAMPUNG	KOTA BANDAR LAMPUNG	1	1
55	RIAU	KAB. BENGKALIS	1	1
56	RIAU	KOTA PEKANBARU	1	0
57	SULAWESI UTARA	KOTA KOTAMOBAGU	1	0
58	SUMATERA SELATAN	KOTA PALEMBANG	2	1
59	SUMATERA SELATAN	KOTA PRABUMULIH	1	1

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
60	SUMATERA UTARA	KOTA MEDAN	7	7

Dari tabel yang ditunjukkan diatas, dapat disimpulkan bahwa hasil pengolahan data terdapat 15 Provinsi dan 60 Kota/Kabupaten yang didapat dari hasil 413 *records* transaksi penjualan Oktober – Desember 2022. Dan terdapat selisih dari jumlah “Total Transaksi” dan “Transaksi Selesai” dikarenakan adanya beberapa transaksi yang dibatalkan. Pembatalan transaksi dapat dipengaruhi oleh berbagai hal, seperti ingin mengubah alamat, memasukkan/mengubah kode *voucher*, mengubah pesaanan, mengubah metode pembayaran, dan sebagainya. Dari data diatas disimpulkan terdapat 413 total transaksi, 350 transaksi selesai dan 63 transaksi yang dibatalkan.

2) Data Preprocessing

Tahap *data preprocessing* yaitu tahap mempersiapkan data mentah yang akan diolah terlebih dahulu. Biasanya dilakukan dengan mengeleminasi data yang kurang sesuai dan kemudian dilakukan pembersihan data seperti *missing value*. *Missing value* dapat terjadi jika informasi untuk suatu objek tidak diberikan, sulit dicari atau memang tidak ada. Namun dalam data penelitian ini tidak memiliki *missing value* sehingga dapat langsung dilanjutkan pada tahap berikutnya.

Automatically detect type

Info
60 instances
4 features (no missing values)
Data has no target variable.
1 meta attribute

Columns (Double click to edit)

Name	Type	Role	Values
1 No.	S text	meta	
2 Provinsi	C categorical	feature	BALI, BANTEN, DI YOGYAKARTA, DKI JAKARTA, JAWA BARAT, JAWA TENGAH, JAWA TIMUR, ...
3 Total Transaksi	N numeric	feature	
4 Transaksi Selesai	N numeric	feature	
5 Kota/Kabupaten	C categorical	feature	

Gambar 4.5 : Data tidak memiliki *missing value*

Dari gambar 4.5 dapat terlihat jika data yang di *input* memiliki 60 baris (*instances*), 4 fitur (*features*) dengan keterangan tidak ada *missing value*, data tidak memiliki *target* dan memiliki 1 *meta* (fitur yang tidak akan digunakan/berperan dalam penelitian).

3) *Data Transformation*

Tahap *Data Transformation* yaitu tahap dimana mengubah data terpilih ke dalam bentuk yang sesuai prosedur *data mining* dimana dengan meringkas atau menggabungkan operasi. Transformasi data ini yaitu mengubah standar angka dalam data menjadi 0-1. Dan teknik yang digunakan adalah *Standard Scaler*, yang bertujuan untuk membuat rata-rata 0 dan variasi 1. Nilai 0 menandakan “Tidak Optimal” sedangkan angka 1 menandakan “Optimal”. Sebagai contoh pada tabel dibawah, Kota Kotamobagu di provinsi Sumatra Utara memiliki nilai total transaksi dan transaksi selesai adalah 0.00, angka tersebut menandakan bahwa tidak ada transaksi yang terjadi pada wilayah tersebut. Sedangkan pada Jakarta Barat di provinsi DKI Jakarta memiliki nilai total transaksi dan transaksi selesai adalah 1, maka menandakan bahwa transaksi penjualan pada wilayah tersebut sudah optimal. Kesimpulannya jika nilai mendekati angka 1, maka transaksi penjualan di wilayah tersebut mendekati optimal. Begitu pula dengan sebaliknya.

Untuk menghitung nilai transformasi data dapat menggunakan rumus sebagai berikut :

$$Z = (X - \mu) / \sigma$$

Keterangan :

Z adalah nilai hasil dari standarisasi.

X adalah nilai asli dari atribut yang akan diskalakan.

μ adalah rata-rata (mean) dari atribut tersebut di seluruh data.

σ adalah standar deviasi (standard deviation) dari atribut tersebut di seluruh data.

Tahap awal adalah menghitung “Total Transaksi” kemudian dibagi dengan banyaknya data (413 transaksi : 60 data wilayah = 6,883) dan hasilnya adalah nilai μ . Setelah itu menghitung standar deviasi yaitu dengan rumus :

$$\sigma = \sqrt{(\sum(x_i - \mu)^2 / N)}$$

Keterangan :

σ adalah standar deviasi.

Σ merupakan simbol sigma yang menunjukkan penjumlahan.

x_i adalah setiap nilai dalam data.

μ adalah rata-rata (mean) dari data.

N adalah jumlah total data.

$$\sigma = \sqrt{[(1-6.883)^2 + (2-6.883)^2 + (9-6.883)^2 + \dots / 60] = ?}$$

Kemudian tahap terakhir

Nilai transformasi (per wilayah) = (Total transaksi (per wilayah) – 6.883 (hasil μ)) / hasil σ .

Tabel 4.2 : Data Setelah Transformasi

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
1	BALI	KAB. GIANYAR	0.00	0.0263
2	BALI	KOTA DENPASAR	0.0222	0.0526
3	BANTEN	KAB. TANGERANG	0.1778	0.2368
4	BANTEN	KOTA CILEGON	0.00	0.0263
5	BANTEN	KOTA SERANG	0.2889	0.3421
6	BANTEN	KOTA TANGERANG	0.2222	0.2895
7	BANTEN	KOTA TANGERANG SELATAN	0.1778	0.1842
8	DI YOGYAKARTA	KAB. BANTUL	0.00	0.0263
9	DI YOGYAKARTA	KAB. KULON PROGO	0.0889	0.0789
10	DI YOGYAKARTA	KAB. SLEMAN	0.0444	0.0789
11	DKI JAKARTA	KOTA JAKARTA BARAT	1.00	1.00
12	DKI JAKARTA	KOTA JAKARTA PUSAT	03556	0.4211
13	DKI JAKARTA	KOTA JAKARTA SELATAN	0.40	0.3947
14	DKI JAKARTA	KOTA JAKARTA TIMUR	0.3778	0.1842
15	DKI JAKARTA	KOTA JAKARTA UTARA	0.4889	0.3421
16	JAWA BARAT	KAB. BANDUNG	0.1333	0.1579
17	JAWA BARAT	KAB. BANDUNG BARAT	0.0222	0.0526
18	JAWA BARAT	KAB. BEKASI	0.2444	0.2895
19	JAWA BARAT	KAB. BOGOR	0.4667	0.4474

No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
20	JAWA BARAT	KAB. CIANJUR	0.00	0.0263
21	JAWA BARAT	KAB. INDRAMAYU	0.00	0.0263
22	JAWA BARAT	KAB. KARAWANG	0.00	0.0263
23	JAWA BARAT	KAB. KUNINGAN	0.00	0.0263
24	JAWA BARAT	KAB. SUBANG	0.0222	0.0526
25	JAWA BARAT	KAB. SUKABUMI	0.00	0.0263
26	JAWA BARAT	KAB. SUMEDANG	0.00	0.0263
27	JAWA BARAT	KAB. TASIKMALAYA	0.00	0.0263
28	JAWA BARAT	KOTA BANDUNG	0.2222	0.2632
29	JAWA BARAT	KOTA BEKASI	0.5111	0.5789
30	JAWA BARAT	KOTA BOGOR	0.1111	0.1316
31	JAWA BARAT	KOTA CIMAHI	0.0222	0.0526
32	JAWA BARAT	KOTA CIREBON	0.0222	0.0526
33	JAWA BARAT	KOTA DEPOK	0.80	0.8421
34	JAWA BARAT	KOTA SUKABUMI	0.0222	0.0526
35	JAWA BARAT	KOTA TASIKMALAYA	0.0444	0.0789
36	JAWA TENGAH	KAB. BOYOLALI	0.00	0.0263
37	JAWA TENGAH	KAB. BREBES	0.00	0.0263
38	JAWA TENGAH	KAB. JEPARA	0.00	0.0263
39	JAWA TENGAH	KAB. KARANGANYAR	0.00	0.0263
40	JAWA TENGAH	KAB. KLATEN	0.0222	0.0526
41	JAWA TENGAH	KAB. PEMALANG	0.1111	0.1579
42	JAWA TENGAH	KAB. SRAGEN	0.00	0.0263
43	JAWA TENGAH	KOTA SEMARANG	0.1111	0.1579

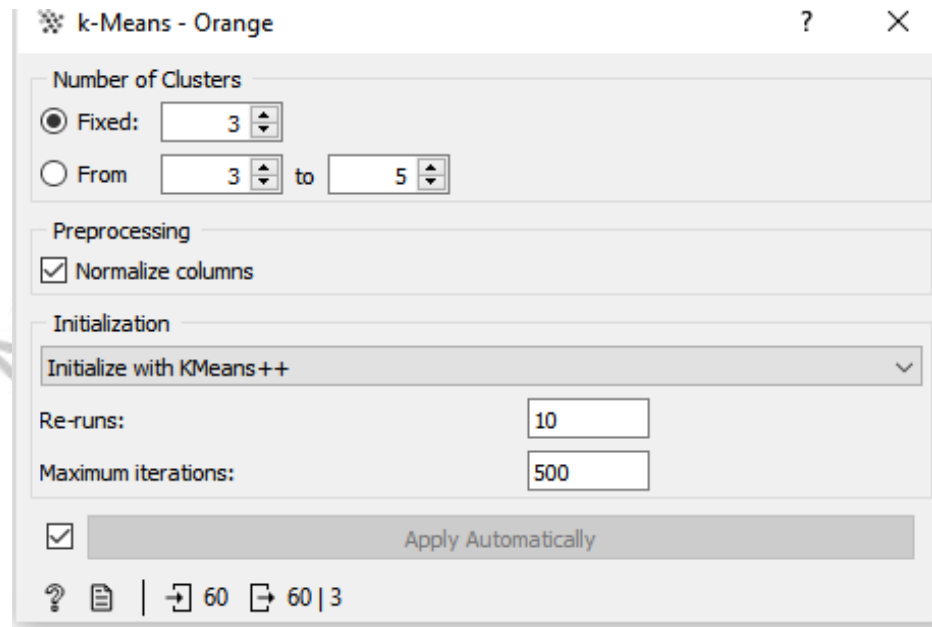
44	JAWA TIMUR	KAB. GRESIK	0.00	0.0263
No.	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai
45	JAWA TIMUR	KAB. MALANG	0.00	0.0263
46	JAWA TIMUR	KAB. SIDOARJO	0.0444	0.0526
47	JAWA TIMUR	KOTA KEDIRI	0.00	0.0263
48	JAWA TIMUR	KOTA MADIUN	0.0222	0.0526
49	JAWA TIMUR	KOTA MALANG	0.1556	0.2105
50	JAWA TIMUR	KOTA SURABAYA	0.6667	0.7105
51	KALIMANTAN BARAT	KOTA PONTIANAK	0.0444	0.0789
52	KALIMANTAN TIMUR	KAB. KUTAI TIMUR	0.0667	0.1053
53	KEPULAUAN RIAU	KOTA BATAM	0.1556	0.2105
54	LAMPUNG	KOTA BANDAR LAMPUNG	0.00	0.0263
55	RIAU	KAB. BENGKALIS	0.00	0.0263
56	RIAU	KOTA PEKANBARU	0.00	0.00
57	SULAWESI UTARA	KOTA KOTAMOBAGU	0.00	0.00
58	SUMATERA SELATAN	KOTA PALEMBANG	0.0222	0.0263
59	SUMATERA SELATAN	KOTA PRABUMULIH	0.00	0.0263
60	SUMATERA UTARA	KOTA MEDAN	0.1333	0.1842

Tabel diatas menunjukkan data yang telah diubah menjadi standar angka atau ditransformasikan hingga sesuai dengan prosedur *data mining*.

4) *Data Mining*

A) *K-Means*

Pada tahap inilah mulai diterapkannya metode *clustering* dari data yang telah ditransformasikan sebelumnya. Saat memilih fitur *K-means*, maka akan muncul gambar seperti dibawah ini :



Gambar 4.6 : *K-means* pada *Orange Data Mining*

Pada gambar 4.6, terdapat menu “Number of Clusters” untuk menentukan jumlah *cluster* (kelompok) yang ingin dibuat. Jika sudah mengetahui berapa *cluster* yang diinginkan, maka dapat memilih “Fixed” dan memasukkan jumlah *cluster*-nya. Jika belum mengetahui ingin membuat berapa *cluster*, maka dapat memilih “From ... to ...” dengan memasukkan jumlah minimal dan maksimal agar sistem dapat memutuskan berapa banyak *cluster* yang dapat dibuat berdasarkan data yang ada namun tidak melebihi minimal dan maksimal *cluster* yang telah ditentukan sebelumnya.

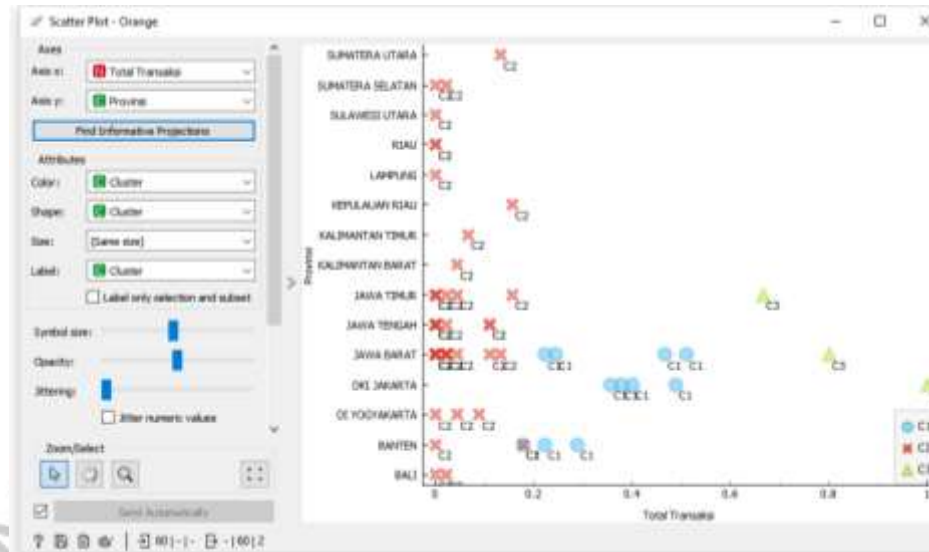
Pada penelitian ini, *cluster* yang ingin dibuat sudah ditentukan sebanyak 3 *cluster* (Penjualan Tertinggi, Penjualan Menengah dan Penjualan Terendah) maka dipilihlah “Fixed” dengan jumlah 3 *cluster*.

Untuk menentukan data masuk ke *cluster* mana, maka digunakannya algoritma *k-means*. Algoritma *k-means* hanya berlaku untuk atribut

numeric dan membutuhkan jumlah *cluster* yang ingin dibuat. Dari data diatas terdapat 2 atribut *numeric*, yaitu “Total Transaksi” dan “Transaksi Selesai”. Jumlah *cluster* yang ingin dibuat pun sudah ditentukan yaitu 3 *cluster*. Pada proses awal, pusat kelompok akan ditentukan secara acak oleh sistem kemudian setiap titik data akan ditandai ke dalam pusat *cluster* terdekat. Posisi pusat *cluster* akan dihitung ulang hingga semua data terklasifikasi ke dalam setiap pusat kelompok dan akhirnya membentuk pusat kelompok baru. Hal tersebut akan terus dilakukan secara berulang hingga anggota dari setiap *cluster* tidak berubah lagi. Kemudian data akan dimasukkan dalam *cluster* terdekat. Metode yang digunakan adalah jarak *Euclidean* yang dapat digunakan untuk menghitung jarak antara titik data dengan pusat *cluster*. Dan jika ada 2 titik data yang memiliki tingkat kedekatan yang tinggi, maka kemungkinan 2 titik data tersebut dikelompokkan pada *cluster* yang sama.

B) *Scatter Plot*

Merupakan sebuah grafik yang menggunakan banyak titik untuk menggambarkan/mewakili suatu nilai dari 2 variabel numerik yang berbeda. Setiap titik memiliki sumbu horizontal dan vertikal untuk menunjukkan nilai data individual dan biasa digunakan untuk mengamati hubungan antar variabel dengan memberikan tampilan secara visual. *Scatter Plot* juga merupakan salah satu fitur yang disediakan pada *Orange Data Mining*. Dan setelah tahap *k-means* maka dilanjutkan dengan menyambungkannya ke *Scatter plot*.



Gambar 4.7 : Scatter Plot

Gambar 4.7 merupakan tampilan *Scatter plot* pada *Orange Data Mining*, Menu “*Axes*” digunakan untuk menentukan sumbu x (horizontal) dan sumbu y (vertikal). Kemudian terdapat menu “*Attributes*” yang dapat digunakan untuk mengatur tampilan dari titik-titik yang ada pada grafik, seperti “*Color*” untuk warna, “*Shape*” untuk bentuk, “*Size*” untuk ukuran dan “*Label*” untuk menambahkan keterangan.

Untuk penelitian ini diatur sebagai berikut :

- Pada menu “*Axes*” :
 - Sumbu x : Total Transaksi
 - Sumbu y : Provinsi
- Pada menu “*Attributes*” :
 - *Color* : Dikelompokkan berdasarkan *Cluster*, sehingga setiap *cluster* memiliki warna yang berbeda.
 - *Shape* : Dikelompokkan berdasarkan *Cluster*, sehingga setiap *cluster* memiliki bentuk yang berbeda.
 - *Size* : Dibuat “*Same Size*” sehingga semua titik memiliki ukuran yang sama.

- *Label* : Dikelompokkan berdasarkan *Cluster*, sehingga setiap titik terdapat keterangan bahwa titik tersebut termasuk dalam kategori *cluster* yang mana.

C) *Table Clustering*

No.	Cluster	Provinsi	Kota/Kabupaten	Transaksi Selesai	Total Transaksi	Silhouette
1	C2	BALI	KAB. GIANYAR	0.00	0.0263	0.602338
2	C2	BALI	KOTA DENPASAR	0.0222	0.0526	0.595953
3	C1	BANTEN	KAB. TANGERANG	0.1778	0.2368	0.506869
4	C2	BANTEN	KOTA CILEGON	0.00	0.0263	0.594651
5	C1	BANTEN	KOTA SERANG	0.2889	0.3421	0.575342
6	C1	BANTEN	KOTA TANGERANG	0.2222	0.2895	0.542953
7	C2	BANTEN	KOTA TANGERANG SELATAN	0.1778	0.1842	0.501426
8	C2	DI YOGYAKARTA	KAB. BANTUL	0.00	0.0263	0.603493

Gambar 4.8 : Tabel *Clustering* pada *Orange Data Mining*

Tabel 4.3 : Tabel *Clustering*

No.	Cluster	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai	Silhouette
1	C2	BALI	KAB. GIANYAR	0.00	0.0263	0.602338
2	C2	BALI	KOTA DENPASAR	0.0222	0.0526	0.595953
3	C1	BANTEN	KAB. TANGERANG	0.1778	0.2368	0.506869
4	C2	BANTEN	KOTA CILEGON	0.00	0.0263	0.594651
5	C1	BANTEN	KOTA SERANG	0.2889	0.3421	0.575342
6	C1	BANTEN	KOTA TANGERANG	0.2222	0.2895	0.542953
7	C2	BANTEN	KOTA TANGERANG SELATAN	0.1778	0.1842	0.501426
8	C2	DI YOGYAKARTA	KAB. BANTUL	0.00	0.0263	0.603493

		TA				
9	C2	DI YOGYAKARTA	KAB. KULON PROGO	0.0889	0.0789	0.580093
No.	Cluster	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai	Silhouette
10	C2	DI YOGYAKARTA	KAB. SLEMAN	0.0444	0.0789	0.589274
11	C3	DKI JAKARTA	KOTA JAKARTA BARAT	1.00	1.00	0.636087
12	C1	DKI JAKARTA	KOTA JAKARTA PUSAT	03556	0.4211	0.604775
13	C1	DKI JAKARTA	KOTA JAKARTA SELATAN	0.40	0.3947	0.607861
14	C1	DKI JAKARTA	KOTA JAKARTA TIMUR	0.3778	0.1842	0.561367
15	C1	DKI JAKARTA	KOTA JAKARTA UTARA	0.4889	0.3421	0.607976
16	C2	JAWA BARAT	KAB. BANDUNG	0.1333	0.1579	0.548293
17	C2	JAWA BARAT	KAB. BANDUNG BARAT	0.0222	0.0526	0.603831
18	C1	JAWA BARAT	KAB. BEKASI	0.2444	0.2895	0.540886
19	C1	JAWA BARAT	KAB. BOGOR	0.4667	0.4474	0.60187
20	C2	JAWA BARAT	KAB. CIANJUR	0.00	0.0263	0.610355
21	C2	JAWA BARAT	KAB. INDRAMAYU	0.00	0.0263	0.610355
22	C2	JAWA BARAT	KAB. KARAWANG	0.00	0.0263	0.610355
23	C2	JAWA BARAT	KAB. KUNINGAN	0.00	0.0263	0.610355
24	C2	JAWA BARAT	KAB. SUBANG	0.0222	0.0526	0.603831
25	C2	JAWA BARAT	KAB. SUKABUMI	0.00	0.0263	0.610355

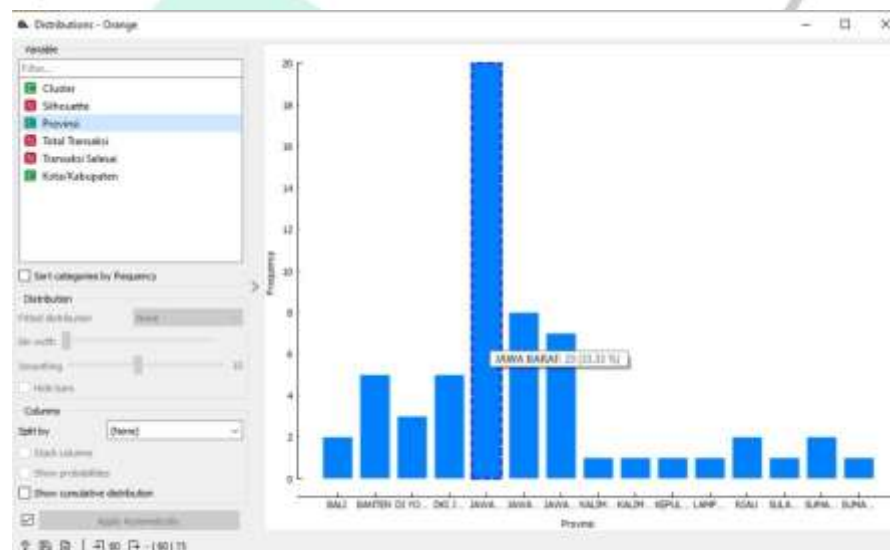
No.	Cluster	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai	Silhouette
26	C2	JAWA BARAT	KAB. SUMEDANG	0.00	0.0263	0.610355
27	C2	JAWA BARAT	KAB. TASIKMALAYA	0.00	0.0263	0.610355
28	C1	JAWA BARAT	KOTA BANDUNG	0.2222	0.2632	0.522285
29	C1	JAWA BARAT	KOTA BEKASI	0.5111	0.5789	0.535135
30	C2	JAWA BARAT	KOTA BOGOR	0.1111	0.1316	0.564862
31	C2	JAWA BARAT	KOTA CIMAHI	0.0222	0.0526	0.603831
32	C2	JAWA BARAT	KOTA CIREBON	0.0222	0.0526	0.603831
33	C3	JAWA BARAT	KOTA DEPOK	0.80	0.8421	0.619266
34	C2	JAWA BARAT	KOTA SUKABUMI	0.0222	0.0526	0.603831
35	C2	JAWA BARAT	KOTA TASIKMALAYA	0.0444	0.0789	0.59535
36	C2	JAWA TENGAH	KAB. BOYOLALI	0.00	0.0263	0.609374
37	C2	JAWA TENGAH	KAB. BREBES	0.00	0.0263	0.609374
38	C2	JAWA TENGAH	KAB. JEPARA	0.00	0.0263	0.609374
39	C2	JAWA TENGAH	KAB. KARANGANYAR	0.00	0.0263	0.609374
40	C2	JAWA TENGAH	KAB. KLATEN	0.0222	0.0526	0.603422
41	C2	JAWA TENGAH	KAB. PEMALANG	0.1111	0.1579	0.561572
42	C2	JAWA TENGAH	KAB. SRAGEN	0.00	0.0263	0.609374
43	C2	JAWA TENGAH	KOTA SEMARANG	0.1111	0.1579	0.561572

No.	Cluster	Provinsi	Kota/Kabupaten	Total Transaksi	Transaksi Selesai	Silhouette
44	C2	JAWA TIMUR	KAB. GRESIK	0.00	0.0263	0.606981
45	C2	JAWA TIMUR	KAB. MALANG	0.00	0.0263	0.606981
46	C2	JAWA TIMUR	KAB. SIDOARJO	0.0444	0.0526	0.597429
47	C2	JAWA TIMUR	KOTA KEDIRI	0.00	0.0263	0.606981
48	C2	JAWA TIMUR	KOTA MADIUN	0.0222	0.0526	0.600888
49	C2	JAWA TIMUR	KOTA MALANG	0.1556	0.2105	0.526074
50	C3	JAWA TIMUR	KOTA SURABAYA	0.6667	0.7105	0.558734
51	C2	KALIMANTAN BARAT	KOTA PONTIANAK	0.0444	0.0789	0.586593
52	C2	KALIMANTAN TIMUR	KAB. KUTAI TIMUR	0.0667	0.1053	0.576743
53	C2	KEPULAUAN RIAU	KOTA BATAM	0.1556	0.2105	0.518756
54	C2	LAMPUNG	KOTA BANDAR LAMPUNG	0.00	0.0263	0.601119
55	C2	RIAU	KAB. BENGKALIS	0.00	0.0263	0.60234
56	C2	RIAU	KOTA PEKANBARU	0.00	0.00	0.605006
57	C2	SULAWESI UTARA	KOTA KOTAMOBAGU	0.00	0.00	0.603817
58	C2	SUMATERA SELATAN	KOTA PALEMBANG	0.0222	0.0263	0.599459
59	C2	SUMATERA SELATAN	KOTA PRABUMULIH	0.00	0.0263	0.602342
60	C2	SUMATERA UTARA	KOTA MEDAN	0.1333	0.1842	0.535904

Pada gambar 4.7 dan tabel 4.3, terdapat 2 kolom tambahan yaitu *cluster* dan *silhouette*. Kolom *cluster* berisikan informasi wilayah tersebut masuk dalam kelompok (*cluster*) yang mana. Sedangkan kolom *silhouette* menunjukkan ukuran seberapa mirip suatu objek dengan *cluster*-nya, nilai *silhouette* ini berkisaran antara -1 hingga 1, semakin tinggi nilai *silhouette* maka menunjukkan semakin tinggi pula tingkat kecocokkannya.

D) Distributions

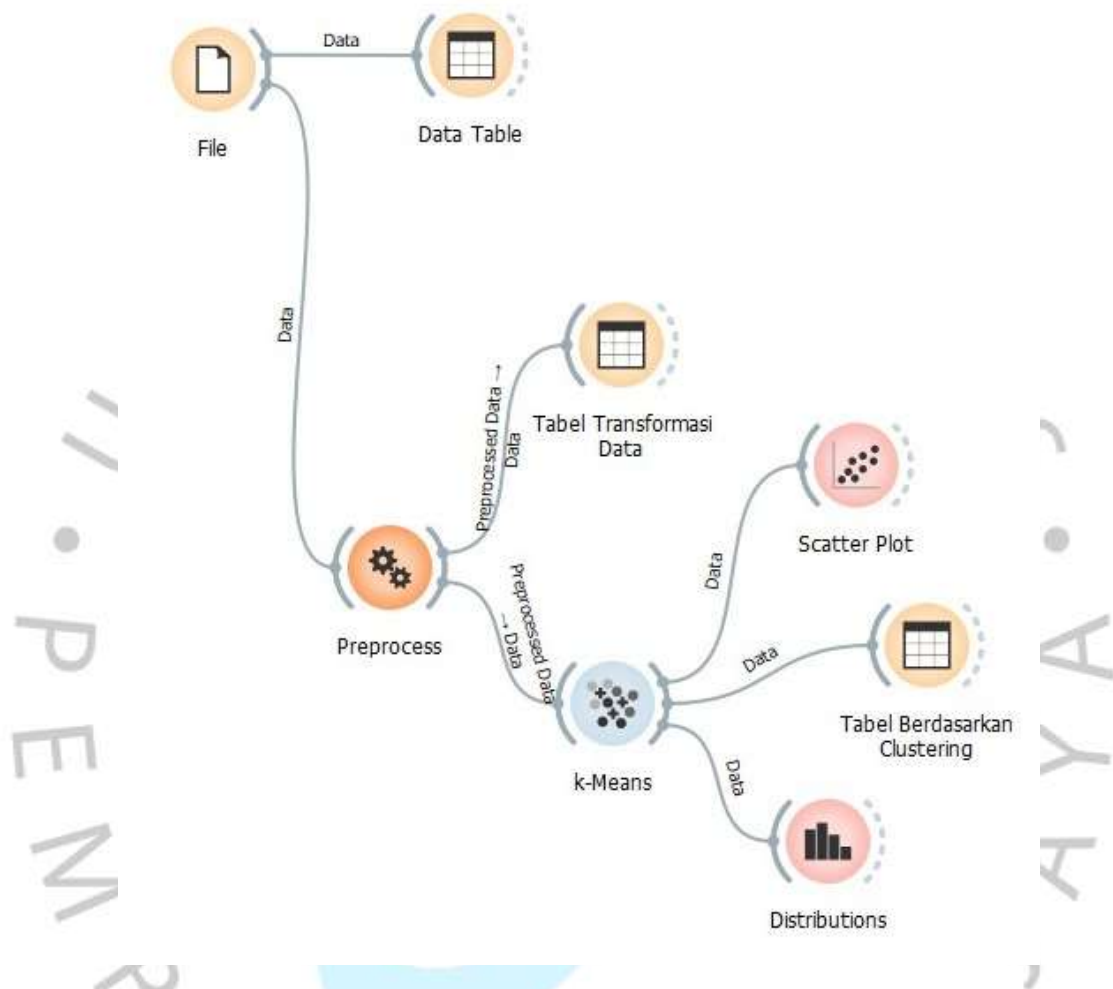
Fitur ini digunakan untuk memvisualisasikan data menjadi diagram batang agar mempermudah dalam membaca data.



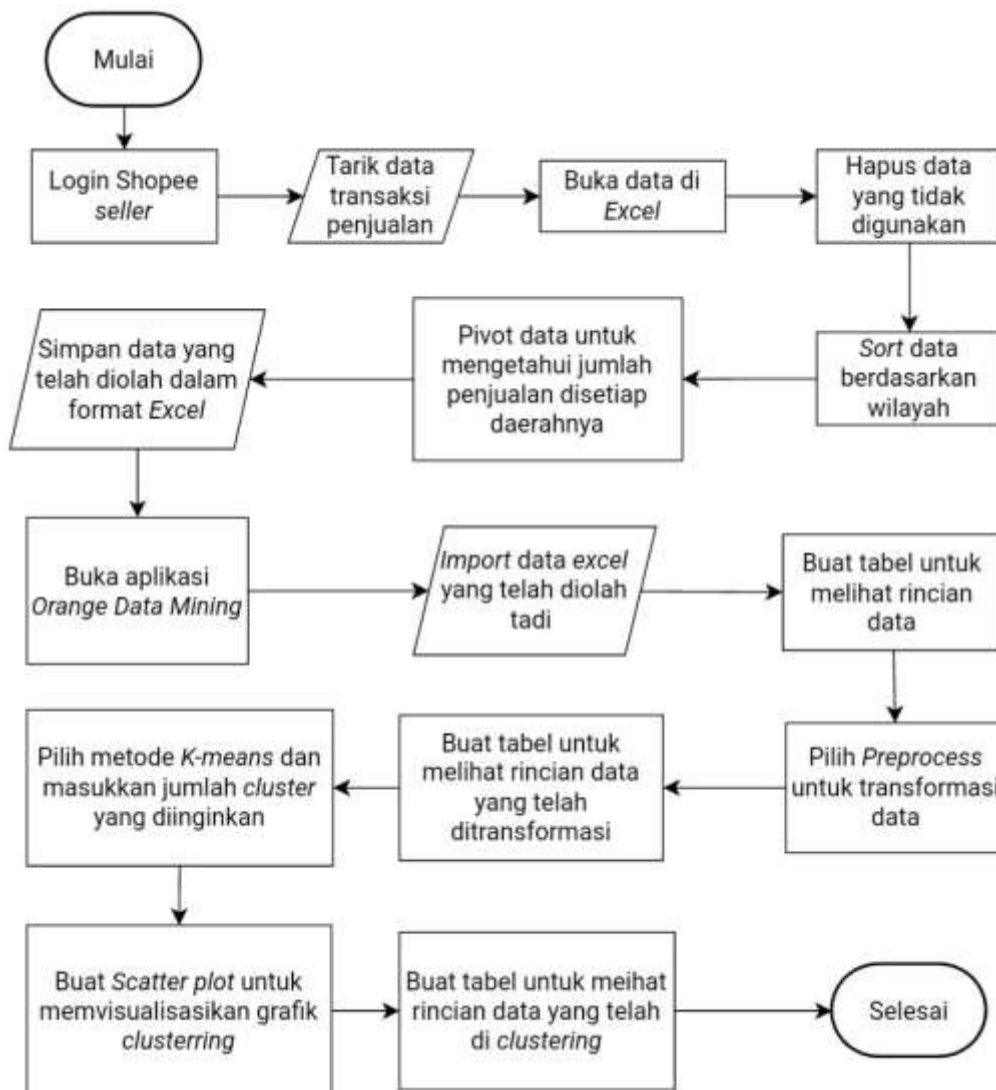
Gambar 4.9 : Distributions pada Orange Data Mining

Pada gambar 4.9 diatas menampilkan diagram batang yang menunjukkan tingkat frekuensi per provinsi. Dapat dilihat Jawa Barat merupakan provinsi dengan frekuensi penjualan tertinggi, hal ini dapat dipengaruhi dari jumlah kota/kabupaten yang berada di Jawa Barat. Pada gambar diatas terdapat *numeric* 20 yang menandakan jumlah kota/kabupaten yang ada di Jawa Barat. Dan persentasi 33.33% menandakan persentasi “Total Transaksi” pada Jawa Barat sebanyak 33,33% dari jumlah semua “Total Transaksi” seluruh wilayah.

Dan jika semua tahap disusun dalam *Orange Data Mining*, kurang lebih akan terlihat seperti gambar dibawah ini :



Gambar 4.10 : Proses *Clustering K-means* pada *Orange Data Mining*



Dari penjelasan diatas, dapat digambarkan menjadi *Flowchart* dibawah ini :

Gambar 4.11 : *Flowchart* Sistem Baru

Dari gambar 3.4 yaitu *Flowchart* Sistem Lama dan gambar 4.11 yaitu *Flowchart* Sistem Baru, dapat dilihat untuk sistem lama tahap dimulai dari menarik data dari Shopee Seller hingga menyimpan data yang telah diolah (Pivot) dalam format Excel. Sedangkan pada sistem baru, data yang telah diolah tersebut kemudian di-import ke aplikasi Orange Data Mining untuk kemudian olah dengan proses metode *k-means* yang merupakan salah satu metode dalam *Data Mining*. Pada sistem baru ini, data akan dibagi menjadi 3 cluster yang dapat membantu

perusahaan untuk membuat strategi bisnis berdasarkan kelompoknya. Dan data dapat divisualisasi dalam bentuk *Scatter plot* dan diagram batang (*Distributions*).

