

BAB IV

HASIL PEMBAHASAN

4.1. Hasil Pengumpulan Data

Pada bab ini akan dibahas mengenai hasil dari pengumpulan data konsumen CV. Lombok Mas Kraksaan yang akan di olah dengan menggunakan *agglomerative hierarchical clustering*.

4.1.1. Observasi

Kegiatan observasi dilakukan pertama kali pada tanggal 20 April 2021. Observasi dilakukan dengan admin marketing dan sales conter Cv. Lombok Mas Kraksaan. Dalam kegiatan ini dilakukan dengan mengamati proses pelayanan konsumen dan pendataan konsumen, proses pelayanan konsumen terkait pembelian unit motor yaitu dengan menjelaskan spesifikasi dan keunggulan motor kepada konsumen, serta proses transaksi yang tersedia. Setelah proses pelayanan maka proses selanjutnya adalah pendataan konsumen

Kegiatan observasi kedua dilakukan pada tanggal 28 April 2021 yaitu mengamati proses pendataan dalam pembelian unit motor, konsumen harus menyerahkan data – data yang ditentukan guna untuk proses pendataan pada konsumen. sehingga mempermudah dalam proses transaksi pembelian.

4.1.2. Wawancara

Pada tahap ini dilakukan dengan cara mengajukan pertanyaan – pertanyaan langsung kepada pihak yang bersangkutan. Wawancara dilakukan pada tanggal 20 April 2021 dengan narasumber admin marketing dan sales conter yang menghasilkan pertanyaan sebagai berikut :

1. Proses pendataan yang masih dilakukan secara manual, masih tidak otomatis.
2. Untuk laporan penjualan dan rencana target penjualan yang masih kurang efektif, sehingga sulit dalam melakukan promosi.
3. Untuk proses penjelasan spesifikasi dan keunggulan motor, sales menjelaskan secara perinci terhadap konsumen dan proses transaksi yang tersedia sehingga mempermudah konsumen dalam melakukan pembelian.

4.2. Analisis Data

Setelah melakukan pengumpulan data, langkah selanjutnya adalah melakukan analisis data dalam penerapan data mining ini menggunakan tahapan Knowledge Discovery in Database (KDD)

4.2.1. Selection

Proses penyeleksian atau pemilihan data untuk memeriksa data yang inkonsisten, membuang duplikasi data serta memperbaiki kesalahan yang ada pada data.

Pada tabel 4.1. adalah data yang belum di proses dan berjumlah 500 data, dapat dijelaskan bahwa setiap proses dapat dipilih sesuai kebutuhan, data yang baru saja dikumpulkan kemungkinan besar memiliki banyak bagian yang tidak relevan. Oleh karna itu perlu adanya proses seleksi data

Tabel 4. 1 Sampel Data awal konsumen yang belum di proses

No	Nama	Alamat	Pekerjaan	J.P em bel ian	Type motor	DP	Angsur an	Harga
1	Amina	Suka pura	Buruh Tani	T	Beat sporty cbs iss	-	-	Rp.18.775.000
2	Sulami	Sentu l	Guru	K	Beat sporty cbs	Rp.3.100.0 00	R p.690.00 0	Rp18.100.000
3	Arjo noriya	Brang gah	Guru	K	Vario 125 cbs	Rp.5.250.0 00	R p.535.00 0	Rp.20.710.000
4	Mahrud in	Ngep ung	Buruh Tani	K	Pcx cbs	Rp.8.500.0 00	R p.968.00 0	Rp.31.500.000
5	Yit salma	Wrin ginan om	IRT	T	Scoopy	-	-	Rp.20.500.000
6	Marhab an	Ngadi sari	Guru	T	Vario 125 cbs	-	-	Rp.21.220.000

Tabel 4. 2 (Lanjutan)

N o	Nama	Ala mat	Peker jaan	J.P em bel ian	Type motor	DP	Angsur an	Harga
7	Siti khotija	Lumb ang	Pegawa i Negri	K	Beat sporty cbs iss	Rp.2.950.0 00	R p.732.00 0	Rp.18.775.000
8	Buana	Maro n	Dokter	T	Scoopy	-	-	Rp.20.500.000
9	Muham mad fahmi	Sentu l	Buruh Pabrik	K	Beat street	Rp.3.250.0 00	R p.623.00 0	Rp.18.600.000
1 0	Nuraini azizah	Suka pura	IRT	K	Vario 125 cbs	Rp.5.500.0 00	R p.728.00 0	Rp.21.220.000

4.2.2. Preprocessing

Pre- processing data adalah hal yang harus dilakukan dalam proses data *mining*, karena tidak semua data atau atribut data dalam data tersebut digunakan dalam proses data *mining*.

Pada proses pre-processing atau pembersihan data untuk memeriksa data yang inkonsisten, membuang duplikasi data serta memperbaiki kesalahan yang ada pada data. Pada data konsumen variabel yang tidak digunakan akan dihapus karna tidak semua data atau variabel yang akan duolah. Jadi hanya berfokus pada 3 variabel. Variabel yang digunakan adalah Pekerjaan, Jenis Pembelian dan Type Motor. Dapat dilihat dalam Tabel 4.2.

Tabel 4. 3 data yang telah di cleaning

PEKERJAAN	JENIS PEMBELIAN	TYPE MOTOR
Buruh Tani	Tunai	Beat Sporty Cbs Iss
Guru	Kredit	Beat Sporty Cbs
Guru	Kredit	Vario 125 Cbs
Buruh Tani	Kredit	Pcx Cbs
Ibu Rumah Tangga	Tunai	Scoopy
Guru	Tunai	Vario 125 Cbs
Pegawai Negri	Kredit	Beat Sporty Cbs Iss
Dokter	Tunai	Scoopy
Buruh Pabrik	Kredit	Beat Street
Ibu Rumah Tangga	Kredit	Vario 125 Cbs

4.2.3. Transformasi

Pada tahap ini data yang telah di cleaning di transformasikan menjadi data yang siap diminingkan. Dengan memilih data sesuai kebutuhan yaitu variabel pekerjaan, jenis pembelian, dan type motor. Kemudian dinisialisasi data yang bertujuan untuk melanjutkan proses pengolahan data ke tingkat selanjutnya. Dapat dilihat pada tabel 4.1.

Tabel 4. 3 data yang telah di Inisialisasi

PEKERJAAN	JENIS PEMBELIAN	TYPE MOTOR
1	2	2
9	1	1
9	1	3
1	1	5
2	2	6
9	2	3
3	1	2
4	2	6
5	1	7
2	1	3

4.2.4. Data Mining

Dari data yang siap diminigkan maka pada tahap ini berfokus pada penentuan pola atau informasi dalam data tersebut dengan menggunakan teknik *Clustering* dengan *Algoritma Agglomerative Hierarchical Clustering*.

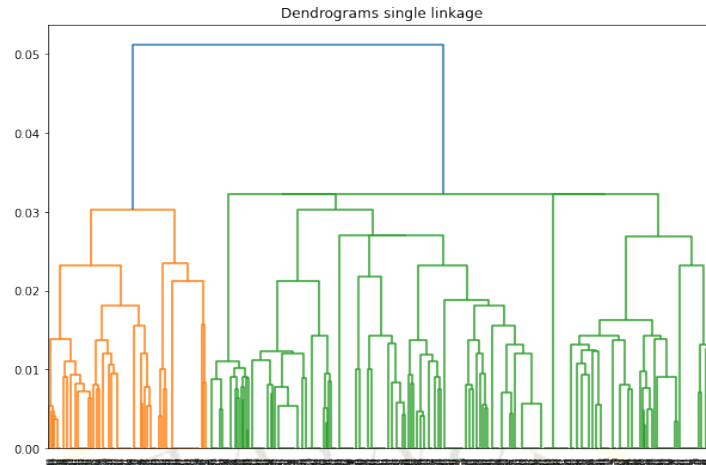
4.2.5. Clustering

Analisis kelompok yang merupakan mengelompokkan suatu data dengan data lainnya yang menggambarkan objek dan hubungan diantara data tersebut. Dengan ini dilakukan seperti diatas supaya data dalam setiap kelompok data yang tercluster yaitu objek yang serupa atau mempunyai similaritas tertinggi. Dalam penelitian ini, pengelompokkan data menggunakan Agglomerative hierarchical clustering (AHC. Penginputan data berupa format pada excel yang diolah menggunakan agglomerative. Data tersebut berupa data penjualan sejumlah 500 dengan 8 atribut diolah dengan menggunakan metode agglomerative, dan menghasilkan kelompok atau cluster.

Dalam penerapan agglomerative terdapat 3 metode, metode tersebut adalah single linkage, complete linkage dan Average Linkage. Dalam perhitungan matrik jarak menggunakan Euclidean distance terhadap 3 metode yang digunakan. Dari proses clustering berupa tabel kelompok, dendogram, dan scatter diagram. Berikut adalah proses pengolahan pada data yang ada dalam penelitian ini.

4.2.6. Single Linkage

Proses pengelompokkan dengan menggunakan metode *single linkage*. Dalam penelitian ini akan mengcluster berdasarkan Jenis Pembelian. Untuk mengcluster jenis pembelian terdapat 2 variabel independent yaitu Pekerjaan dan Type motor (X) dan 1 variabel depedent yaitu jenis pembelian (Y) . cluster dalam dendogram *single linkage*. Sumbu X = Sampel data, Sumbu Y = Jarak. Hasil dendogram *single linkage* terdapat pada gambar 4.3. menampilkan *cluster* (0, 1)



Gambar 4. 4 Hasil Dendogram single linkage dengan 3 cluster

Dapat dilihat pada gambar 4.2. dendogram *single linkage* menunjukkan Jumlah data pada setiap *cluster* tidak seimbang atau memiliki selisih jumlah data yang jauh pada setiap cluster. Jumlah data terkecil pada *cluster* 2 dengan angka 0 dan tertinggi pada *cluster* 1 dengan angka 9.

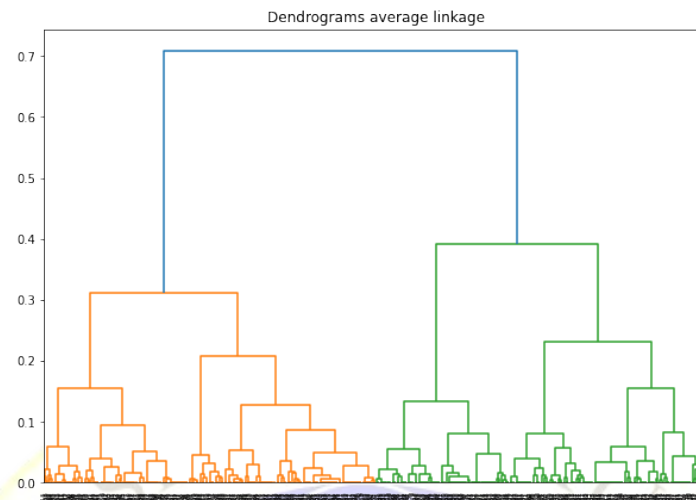
Berikut ini *source code* dalam perhitungan metode *single linkage*:

Segmen Program 4. 1 *dendogram single linkage*

1. `import scipy.cluster.hierarchy as shc`
2. `plt.figure(figsize=(10, 7))`
3. `plt.title("Dendrograms Single linkage")`
4. `plt.xlabel('PEKERJAAN')`
5. `plt.ylabel('TYPE MOTOR')`
6. `dend = shc.dendrogram(shc.linkage(data_scaled, method='single'))`

Setelah proses dendogram selesai menghasilkan 2 cluster yaitu (0, 1). langkah selanjutnya adalah mengelompokkan hierarki untuk 2 cluster dari dendogram di atas. Pertama menggunakan metode *single linkage*. Berikut hasil hirarki pada gambar 4.4.

4.2.7. Average Linkage



Gambar 4. 6 Hasil Dendrogram *average linkage* dengan 3 cluster

Dapat dilihat pada Gambar 4.5. dendrogram *average linkage* lebih seimbang dibandingkan *single linkage*. Jumlah data pada *cluster 0* dan *1* adalah 250. Dan dapat dilihat pada tabel. 4.3

Berikut ini *source code* yang digunakan dalam perhitungan metode *average linkage*:

Segmen Program 4. 2 dendrogram *average linkage*

1. `import scipy.cluster.hierarchy as shc`
2. `plt.figure(figsize=(10, 7))`
3. `plt.title("Dendrograms Average linkage")`
4. `plt.xlabel('PEKERJAAN')`
5. `plt.ylabel('TYPE MOTOR')`
6. `dend = shc.dendrogram(shc.linkage(data_scaled, method='average'))`

Setelah proses dendrogram *average linkage* maka kemudian mengelompokkan hierarki untuk 2 cluster dari dendrogram diatas menggunakan metode *average linkage*:


```
array([[0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0,
1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1,
0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0,
0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0,
0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0,
1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1,
0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1,
1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1,
0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,
0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1,
0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1,
1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,
1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1,
1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0,
1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0,
1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0,
1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1,
1, 1, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1])
```

Gambar 4. 7 hierarki 2 cluster menggunakan Average Linkage

Berikut ini *source code* yang digunakan dalam perhitungan metode *average linkage* :

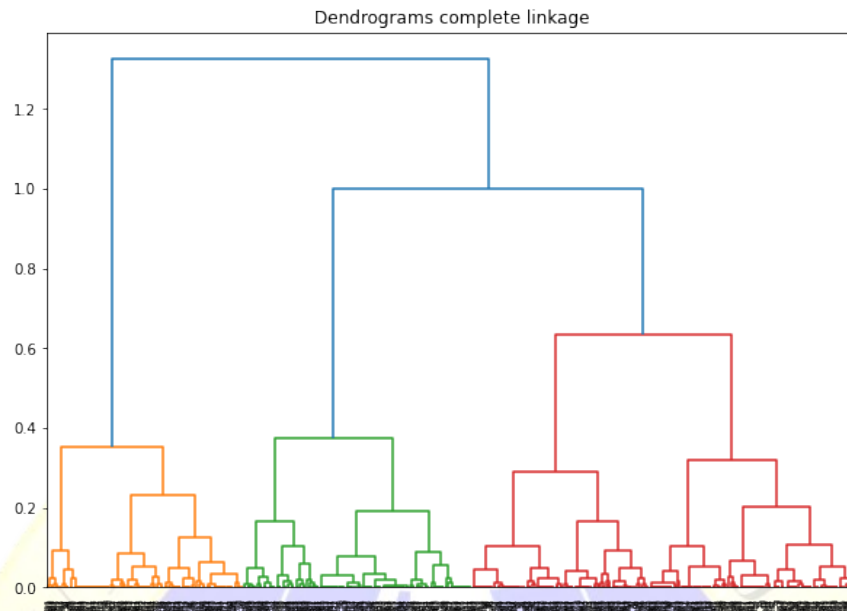
Segmen Program 4. 3 Penerapan hierarki 2 cluster Average Linkage

```
7. from sklearn.cluster import AgglomerativeClustering
8. import numpy as np
9. clusteraverage = AgglomerativeClustering(n_clusters=2, affinity
    = 'euclidean', linkage='average')
10.ca = clusterward.fit_predict(data_scaled)
11.ca
12.xlcs = pd.DataFrame(cs)
13.xlcs.to_excel('hasilclusteraverage.xlsx')
```

Tabel 4. 2 Jumlah data pada setiap Cluster

CLUSTER	JUMLAH
0	250
1	250

4.2.8. Complete Linkage



Gambar 4. 8 Hasil Dendrogram complete linkage dengan 3 cluster

Dapat dilihat pada gambar 4.7. hasil dendrogram pada pengelompokan menggunakan metode *complete linkage* dapat dilihat cluster yang paling sedikit terdapat pada dendrogram berwarna merah. Namun pada metode *complete linkage* terdapat jumlah data pada masing – masing *cluster* masih tidak seimbang dibanding menggunakan metode *average linkage*. Cluster 1 memiliki jumlah data terendah yaitu 121, dan cluster 0 yaitu 379.

Berikut adalah *source code* dalam perhitungan metode *complete linkage*:

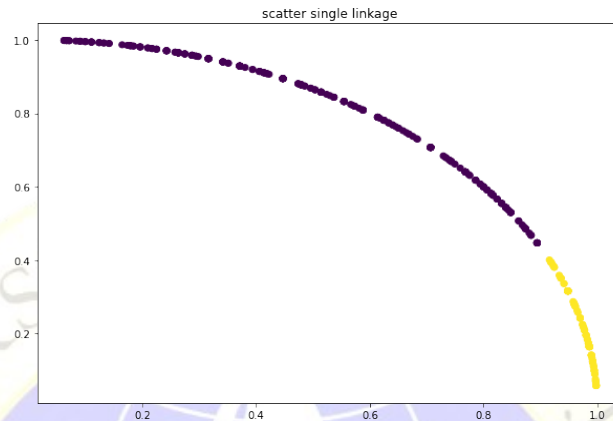
Segmen Program 4. 4 Penerapan dendrogram complete linkage

1. `import scipy.cluster.hierarchy as shc`
2. `plt.figure(figsize=(10, 7))`
3. `plt.title("Dendrograms Complete linkage")`
4. `plt.xlabel('PEKERJAAN')`
5. `plt.ylabel('TYPE MOTOR')`
6. `dend = shc.dendrogram(shc.linkage(data_scaled, method='complete'))`

Setelah proses dendrogram *average linkage* maka kemudian mengelompokkan hierarki untuk 2 cluster dari dendrogram diatas menggunakan metode *complete linkage*:

Dalam penelitian ini terdapat beberapa *scatter diagram* diantaranya adalah *Scatter single linkage*, *scatter average linkage*, dan *scatter complete linkage*.

4.3.1. Scatter Single Linkage



Gambar 4. 10 Scatter Diagram single linkage

Dalam proses *scatter* untuk menentukan Jenis Pembelian dengan 2 titik variabel yaitu (X) Pekerjaan dan (Y) Type Motor. Dapat menunjukkan scatter diagram antara sumbu X dan Y pada gambar 4.9. dapat dilihat antar jarak pada variabel X dan Y. dari bentuk grafik yang dihasilkan, maka grafik dari *scatter Diagram* diatas dinyatakan memiliki hubungan korelasi yang positif. Dimana nilai nilai besar dari variabel X berhubungan dengan nilai besar Variabel Y. sedangkan nilai – nilai kecil variabel X berhubungan dengan nilai Variabel Y

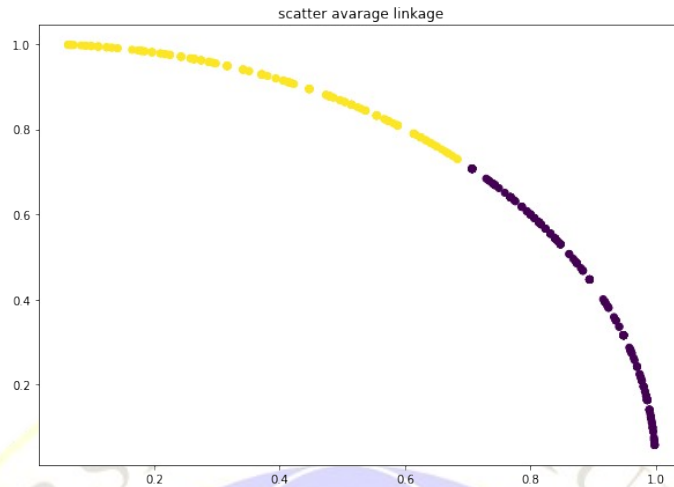
Berikut adalah *source code* yang dalam penerapan *scatter single linkage*:

Segmen Program 4. 6 Penerapan scatter single linkage

1. `import sklearn.metrics as sm`
2. `plt.figure(figsize=(10, 7))`
3. `plt.scatter(data_scaled['TYPE MOTOR'], data_scaled['PEKERJAAN'], c=clustersingle.labels_)`
4. `plt.title('scatter single linkage')`

Hasil Akurasi yang diperoleh dari perhitungan menggunakan metode *single linkage* adalah 15 %

4.3.2. Scatter Average Linkage



Gambar 4. 11 Scatter Diagram average linkage

Dalam proses *scatter* untuk menentukan Jenis Pembelian dengan 2 titik variabel yaitu (X) Pekerjaan dan (Y) Type Motor. Dapat menunjukkan scatter diagram antara sumbu X dan Y pada gambar 4.10. dapat dilihat anatr jarak variabel X dan Y. dari bentuk grafik yang dihasilkan, maka grafik dari *scatter diagram* diatas dinyatakn memiliki hubungan korelasi Negatif. Dimana nila besar dan kecil pada masing – masing variabel X dan Y saling berhubungan.

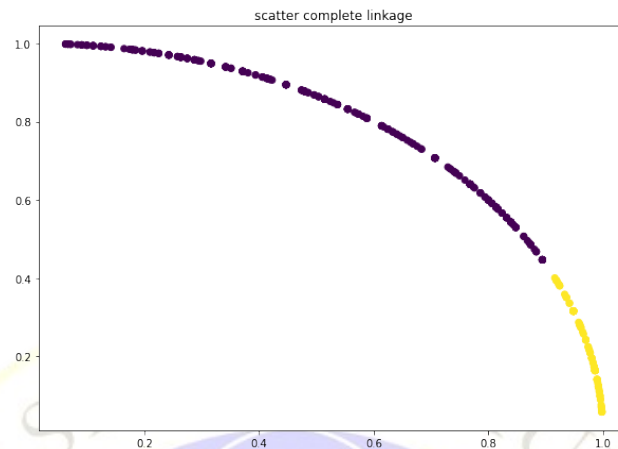
Berikut ini *source code* dalam penerapan *scatter average linkage* :

Segmen Program 4. 7 Penerapan scatter average linkage

1. `import sklearn.metrics as sm`
2. `plt.figure(figsize=(10, 7))`
3. `plt.scatter(data_scaled['TYPE MOTOR'], data_scaled['PEKERJAAN'], c=clust_ersingle.labels_)`
4. `plt.title('scatter average linkage')`

Hasil Akurasi yang diperoleh dari perhitungan menggunakan metode *average linkage* adalah 29 %

4.3.5. Scatter Complete Linkage



Gambar 4. 12 Scatter Diagram *complete linkage*

Dalam proses *scatter* untuk menentukan Jenis Pembelian dengan 2 titik variabel yaitu (X) Type Motor dan (Y) Pekerjaan. Dapat menunjukkan *scatter diagram* antara sumbu X dan Y pada gambar 4.11. dapat dilihat antar jarak pada variabel X dan Y. dari bentuk grafik yang dihasilkan, maka grafik dari *scatter Diagram* diatas dinyatakan memiliki hubungan korelasi yang positif. Dimana nilai nilai besar dari variabel X berhubungan dengan nilai besar Variabel Y. sedangkan nilai – nilai kecil variabel X berhubungan dengan nilai Variabel Y

Berikut adalah *source code* dalam penerapan *scatter complete linkage*:

Segmen Program 4. 8 Penerapan *scatter complete linkage*

1. `import sklearn.metrics as sm`
2. `plt.figure(figsize=(10, 7))`
3. `plt.scatter(data_scaled['TYPE MOTOR'], data_scaled['PEKERJAAN'], c=clustersingle.labels_)`
4. `plt.title('scatter complete linkage')`

Hasil Akurasi yang diperoleh dari perhitungan menggunakan metode *complete linkage* adalah 15 %.

4.4. Perhitungan Manual dengan Matrik Euclidean Distance

Digunakan 10 sampel data dalam perhitungan manual menggunakan matrik Euclidean distance. Jika diketahui matrik jarak antara 10 sampel data yang disajikan pada tabel 4.5.

Tabel 4. 4 Jarak antar objek Cluster 1

	0	1	2	3	4	5	6	7	8	9
0	0	65	65	9	17	65	4	25	41	2
1	65	0	4	80	74	4	37	50	52	53
2	65	4	0	68	58	0	37	34	32	49
3	9	80	68	0	2	68	13	10	20	5
4	17	74	58	2	0	58	17	4	10	9
5	65	4	0	68	58	0	37	34	32	49
6	4	37	37	13	17	37	0	17	29	2
7	25	50	34	10	4	34	17	0	2	13
8	41	52	32	20	10	32	29	2	0	25
9	2	53	49	5	9	49	2	13	25	0

Langkah – langkah dalam algoritma AHC secara umum sebagai berikut:

Mencari objek dengan jarak minimum yaitu 0, maka objek 2 dan 5 bergabung menjadi satu cluster (2,5).

Menghitung jarak antara cluster 2,5 dengan objek lainnya Pada tabel 4.6 berikut menunjukkan bahwa memiliki MIN = 0 adalah (2,5) untuk cluster 2.

Tabel 4. 5 Jarak antar objek Cluster 2

	0	1	2,5	3	4	6	7	8	9
0	0	65	65	9	17	4	25	41	2
1	65	0	4	80	74	37	50	52	53
2,5	65	4	0	68	58	37	34	32	49
3	9	80	68	0	2	13	10	20	5
4	17	74	58	2	0	17	4	10	9
6	4	37	37	13	17	0	17	29	2
7	25	50	34	10	4	17	0	2	13
8	41	52	32	20	10	29	2	0	25
9	2	53	49	5	9	2	13	25	0

Dengan demikian terbentuk matrik jarak yang disajikan pada tabel 4.7. Dan menghasilkan MIN = 2 adalah (0, 6, 7, 8, 9), (3,4). Dan (2,5) .

Tabel 4. 6 Hasil perhitungan jarak

	0, 6, 7, 8, 9	3,4	
0, 6, 7, 8, 9	0		
3,4	4	0	
2,5	32	58	0

Dari *cluster* diatas dapat menghasilkan matrik jarak baru yaitu MIN = 4 dengan hasil cluster terendah yaitu:

- a. Cluster 1 = 0, 3, 4, 6, 7, 8, 9
- b. Cluster 2 = 2, 5

4.5. Analisa Hasil

Merupakan pola informasi yang dihasilkan dari proses data mining. dari hasil tahap – tahap dalam mengolah data konsumen bahwa Implementasi yang telah dijelaskan dapat membantu menganalisis terhadap pengelompokkan pada data konsumen di Cv. Lombok Mas Kraksaan. Dalam proses olah data terhadap 500 konsumen periode Januari – Desember tahun 2020. Dengan demikian dapat diketahui kelompok konsumen berdasarkan Jenis Pembelian.

Dalam pengelompokkan ini, varaibel yang digunakan dalam pengelompokkan yaitu Pekerjaan. Jenis Pembelian dan Type Motor untuk menentukan yang meliputi data konsumen pada bulan Januari – Desember tersebut telah diolah dengan menggunakan *agglomerative hierarchical clustering* (AHC) yang meliputi metode *single linkage*, *average linkage* dan *complete linkage*. Dan pada pengelompokkan hierarki 2 cluser dalam penentuan berdasarkan jenis pembelian sebagai atribut yang digunakan adalah Pekerjaan dan Type Motor yang menghasilkan Cluster = Menghasilkan akurasi yaitu *single linkage* = 15 %, *average linkage* = 29% dan *complete linkage* = 15 %. Perbandingan akurasi tertinggi dari ketiga metode tersebut menggunakan metode *average linkage* = 29%.

Berikut *source code* yang digunakan dalam akurasi pada ketiga metode diatas:

Segmen Program 4.9 proses akurasi

```
1. import sklearn.metrics as sm
```


2. `print('Hasil akurasi dari Single : ', sm.accuracy_score(Y,cw)*100, '%')`
3. `print('Hasil akurasi dari Average : ', sm.accuracy_score(Y,cw)*100, '%')`
4. `print('Hasil akurasi dari Complete : ', sm.accuracy_score(Y,cw)*100, '%')`

a. Dataset Pekerjaan

Hasil perhitungan pada data Pekerjaan total data yang diperoleh sebanyak 500 data dan yang paling banyak yaitu 86 data Pekerjaan, Adapun data pekerjaan paling sedikit yaitu 1 data. Pada gambar dibawah ini terdapat dua baris angka yaitu baris pertama adalah inisialisasi data, baris kedua adalah hasil data yang diperoleh

```

11 86
5 77
1 69
3 51
9 47
12 32
10 30
2 23
13 18
15 15
7 14
16 12
14 11
8 8
17 2
6 2
4 2
18 1
Name: PEKERJAAN, dtype: int64

```

Gambar 4.12 Hasil dataset Pekerjaan

Tabel.4.8 Dataset Pekerjaan

No	Inisialisas i	Pekerjaan	Hasil data
1	11	Pegawai Swasta	86 data
2	5	Buruh Pabrik	77 data
3	1	Buruh Tani	69 data
4	3	Pegawai Negri	51 data
5	9	Guru	47 data

Tabel.4.8 Lanjutan

No	Inisialisas i	Pekerjaan	Hasil data
6	12	Nelayan	32 data
7	10	Pedagang	30 data
8	2	Ibu Rumah Tangga	23 data
9	13	Wiraswasta	18 data
10	15	Ojek	15 data
11	7	Polri	14 data
12	16	Sopir	12 data
13	14	Tukang Becak	11 data
14	8	TNI	8 data
15	17	Pengacara	2 data
16	6	Sales	2 data
17	4	Dokter	2 data
18	18	Wartawan	1 data

b. Dataset Type Motor

Hasil perhitungan pada data Type Motor total data yang diperoleh sebanyak 500 data dan yang paling banyak yaitu 83 data Type Motor Adapun yang paling sedikit yaitu 6 data. Pada gambar dibawah ini terdapat dua baris angka yaitu baris pertama adalah inisialisasi data, baris kedua adalah hasil data yang diperoleh

```

6      83
3      77
4      47
7      44
17     38
1      37
8      33
2      26
9      19
5      17
11     16
14     16
16     13
15     12
10     10
12      6
13      6
Name: TYPE MOTOR, dtype: int64

```

Gambar 4.12 Hasil dataset Type Motor

Tabel 4.9 Dataset Type Motor

No	Inisialisasi	Type Motor	Hasil data
1	6	Scoopy	83 data
2	3	Vario 125 CBS	77 data
3	4	Vario 150 EXC	47 data
4	7	Beat Street	44 data
5	17	Genio	38 data
6	1	Beat Sporty CBS	37 data
7	8	CRF 150	33 data
8	2	Beat Sporty CBS ISS	26 data
9	9	Supra X CW	19 data
10	5	PCX CBS	17 data
11	11	CB 150 STD	16 data
12	14	Revo Fit	16 data
13	16	Revo X	13 data
14	15	ADV CBS	12 data
15	10	CB Verza	10 data
16	12	CBR Merah	6 data
17	13	CBR Hitam	6 data

Tabel 4. 10 Keseimbangan Dendogram

Metode	Dendogram	Jumlah objek percluster		Akurasi
		1	2	
Single Linkage	Tidak	379	131	15 %
	Seimbang			
Average Linkage	Seimbang	250	250	29 %
Complete Linkage	Tidak	379	131	15%
	Seimbang			

4.6. Perbandingan Hasil

Perbandingan hasil dalam penelitian ini adalah perbandingan antara penggunaan clustering secara sistem dan clustering secara manual. Dari tabel diatas Cluster 2 adalah 2,5 dan cluster 1 adalah 0,3,4,6,7,8,9 dengan akurasi data yaitu *single linkage*: 15%, *average linkage*: 29%, dan *complete linkage* : 15%. terdapat beberapa data yang tidak sesuai antara perhitungan sistem dan manual.

Dalam hal ini perbedaan terletak pada penempatan data pada setiap cluster tidak sesuai antara sistem dan manual dikarenakan dalam proses penerapan metode pengelompokkan yang sangat berbeda. Pengelompokkan secara manual hanya menggunakan keseluruhan data pada konsumen yang di cluster, sedangkan sistem menggunakan metode *Euclidean distance* yang kemudian dikelompokkan dengan menggunakan *agglomerative*.

