

Segmentation of Avocado Seed Buyers Using K-Means Clustering

Segmentasi Pembeli Bibit Alpokat Menggunakan K-Means Clustering

Elok Eka Kartika Putri Lestari¹, Daniel Swanjaya², Risky Aswi Ramadhani³

Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: *¹elokekakartika123@gmail.com, ²daniel@unpkediri.ac.id,

³riskyaswiramadhani@gmail.com

Abstract – Al-Fukaat Islamic Boarding School implements retail and wholesale sales in selling avocado seeds. The Al-Fukaat Islamic Boarding School in marketing avocado seeds is still uneven in the marketing areas, namely Java and Bali and the lack of precise targeting of buyers based on regional conditions that are suitable for planting avocado seeds so that sales are not maximized so there are still types of seeds with very little sales. In overcoming this problem, research was carried out using the K-Means Clustering method to obtain buyer segmentation results at the Al-Fukaat Islamic Boarding School. Success in research is driven by SSE (Sum of Squared Error) results. By obtaining the final SSE results, namely 39.7852 applications can be used because they are able to obtain lower SSE results.

Keywords — *k-means, segmentation, SSE*

Abstrak – Pesantren Al-Fukaat menerapkan penjualan secara ecer dan grosir dalam melakukan penjualan bibit alpokat. Pesantren Al-Fukaat dalam memasarkan bibit alpokat masih belum merata pada daerah pemasaran yaitu Jawa dan Bali serta kurang tepatnya target pembeli berdasarkan kondisi wilayah yang cocok untuk penanaman bibit alpokat sehingga penjualan belum maksimal sehingga masih ada jenis bibit dengan penjualan yang sangat sedikit. Dalam mengatasi permasalahan tersebut dilakukan penelitian menggunakan metode *K-Means Clustering* untuk mendapatkan hasil segmentasi pembeli pada Pesantren Al-Fukaat. Keberhasilan dalam penelitian berpacu pada hasil *SSE (Sum of Squared Error)*. Dengan didapatkan hasil akhir *SSE* yaitu 39,7852 aplikasi dapat digunakan karena mampu memperoleh hasil *SSE* yang lebih rendah.

Kata Kunci — *k-means, segmentasi, SSE*

1. PENDAHULUAN

Pesantren Al-Fukaat merupakan tempat pembibitan alpokat yang terletak di Desa Trayang, Kecamatan Ngronggot, Kabupaten Nganjuk. Dalam menjalankan usaha pembibitan alpokat selama 4 tahun Pesantren Al-Fukaat sudah memiliki beberapa mitra bisnis di wilayah Jawa dan Bali. Pesantren ini menjual banyak jenis bibit alpokat import yang diantaranya *Cuba, Bo Khong Ten, Yellow, Red Vietnam* dan *Buccaneer*. Untuk jenis bibit lokal yang dijual di Pesantren ini ada *Alligator, Miki, Hass, Pangeran* dan *Markus*. Pesantren Al-Fukaat menerapkan penjualan secara ecer dan grosir dalam melakukan penjualan bibit alpokat. Pesantren Al-Fukaat dalam memasarkan bibit alpokat masih belum merata pada daerah pemasaran yaitu Jawa dan Bali serta kurang tepatnya target pembeli berdasarkan kondisi wilayah yang cocok untuk penanaman bibit alpokat sehingga penjualan belum maksimal sehingga masih ada jenis bibit dengan penjualan yang sangat sedikit.

Pemasaran bibit alpokat mempunyai peran penting untuk meningkatkan hasil usaha. Pemasaran merupakan proses sosial dan manajerial yang membuat perorangan maupun kelompok memperoleh hal mereka butuhkan dan inginkan dengan menawarkan produk atau jasa dari produsen

hingga konsumen [1]. Pemasaran bibit alpokat juga membutuhkan penyusunan strategi untuk meningkatkan hasil usaha dengan mengumpulkan data dan identifikasi terhadap pelanggan sangat penting dilakukan. Hal tersebut bertujuan untuk memperoleh informasi pelanggan mengenai kebutuhan, gaya hidup, dan karakteristik pelanggan [2].

Maka dalam hal tersebut perlu adanya segmentasi pembeli untuk memaksimalkan penjualan karena adanya strategi pemasaran yang efektif. Segmentasi merupakan sebuah cara untuk membedakan dan mengelompokkan pelanggan menurut golongan [3]. Dengan teknologi saat ini sangat mudah untuk melakukan segmentasi pembeli dengan jumlah yang banyak. Namun dalam menerapkan hal tersebut perlu adanya riset dan survey yang sesuai. Peneliti menggunakan metode *K-Means Clustering* karena metode tersebut mampu mengelompokkan data besar dengan sangat cepat sehingga memudahkan pemilik dalam melakukan segmentasi pembeli [4]. Dalam penelitian ini data yang digunakan adalah data pelanggan pada Pesantren Al-Fukaat sejak tahun 2019 sampai 2022. Cara segmentasi yang dilakukan adalah mengelompokkan data pelanggan sesuai karakteristik, gaya hidup, wilayah dan jumlah bibit yang telah dibeli dengan tahun yang sudah ditentukan.

Penelitian mengenai segmentasi pembeli pernah dilakukan oleh Aditya Pradana, Agus Sopiyan, Tomi Anugrah Pratama, Abdul Yusuf (2022) dengan judul “Analisis strategi pemasaran GoPay berupa segmentasi, targeting, dan positioning serta pengaruhnya terhadap keputusan pembelian” penelitian tersebut menggunakan metode analisis *linier regresi* berganda. Dalam penelitian tersebut hasil yang dapat diterima oleh pengaruh terhadap keputusan pembelian GoPay di Karawang melalui strategi Targeting H2 X2 dan Positioning H3 X3. Dalam penggunaan cara segmenting yang diterapkan masih ada hasil yang negatif (-3.391) [5]. Penelitian selanjutnya berjudul “Perancangan Business Intelligence Dan Segmentasi Pelanggan Menggunakan *K Means Clustering* Berdasarkan *Rfm Model*” tahun 2023 oleh Ahmad Karami dan Yusuf Widharto. Penelitian tersebut menghasilkan kesimpulan bahwa segmentasi pelanggan menggunakan metode *Rfm* dan *K-Means Clustering* dapat dilakukan dan hasil tersebut dapat menggolongkan data berdasarkan skor kriteria pada *RFM*. Permasalahan dalam metode tersebut masih menggunakan *RFM* bukan sepenuhnya menggunakan *K-Means Clustering* [6]. Penelitian selanjutnya berjudul “Implementasi Algoritma *K-Medoids Clustering* Untuk Menentukan Segmentasi Pelanggan” tahun 2023 oleh Nita Mirantika, Tri Septiar Syamfithriani, dan Ragel Trisudarmo penelitian tersebut menggunakan metode *K-Medoids*, analisis *RFM* untuk segmentasi pelanggan pada *XYZ Store* dan *Davies-Bouldin Index (DBI)* dalam mengevaluasi hasil akhir *cluster*. Hasil dalam penelitian tersebut memperoleh 3 segmentasi pelanggan sehingga dapat digunakan untuk segmentasi pelanggan namun penelitian tersebut masih belum menggunakan satu metode saja namun masih menggunakan bantuan analisis *RFM* dan *DBI* dalam mencari hasil segmentasi [7]. Berikutnya ada penelitian yang berjudul “Analisis Segmentasi Pasar Terhadap Volume Penjualan Melalui Keputusan Pembelian (Survei Pada Toko Youth Sneakers)” tahun 2022 oleh Silvia Oktapiani Hayati, menggunakan metode kausalitas dengan pendekatan survei. Dari penelitian tersebut menghasilkan kesimpulan ada pengaruh mediasi Keputusan Pembelian dalam memediasi hubungan Segmentasi Pasar terhadap Volume Penjualan Pada Konsumen Toko Youth Sneakers namun angka mediasi masih 0,075 [8].

Berdasarkan permasalahan diatas maka penulis membuat penelitian berjudul Segmentasi Pembeli Bibit Alpokat menggunakan *K-Means Clustering* dalam penelitian ini. Dalam penelitian ini diharapkan dapat menyelesaikan permasalahan pada Pesantren Al-Fukaat dalam menentukan target pembeli.

2. METODE PENELITIAN

2.1. Metode Penelitian

2.1.1. *K-Means*

Metode *K-Means Clustering* merupakan metode untuk mengelompokkan sebuah objek yang mempunyai kemiripan karakteristik. Jumlah *cluster* yang diinginkan dimaksud dengan *K* sedangkan rata-rata dalam *cluster* data [9].

- 1) Cara perhitungan algoritma *K-Means* dapat dilihat pada persamaan (1) :



$$V = \frac{\sum_{i=1}^n x_i}{n}; i = 1,2,3 \dots \dots n \dots \dots \dots (1)$$

Keterangan :

V : Centroid yang terdapat pada cluster

x_i : data cluster ke-i

n : jumlah data yang menjadi anggota cluster

- a. Menentukan cluster (k) yang akan dibentuk
- b. Menentukan nilai k centroid secara acak/random.
- c. Menghitung jarak data cluster dengan centroid menggunakan persamaan (2) :

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1,2,3 \dots \dots n \dots \dots \dots (2)$$

Keterangan :

x_i : banyaknya data x ke-i

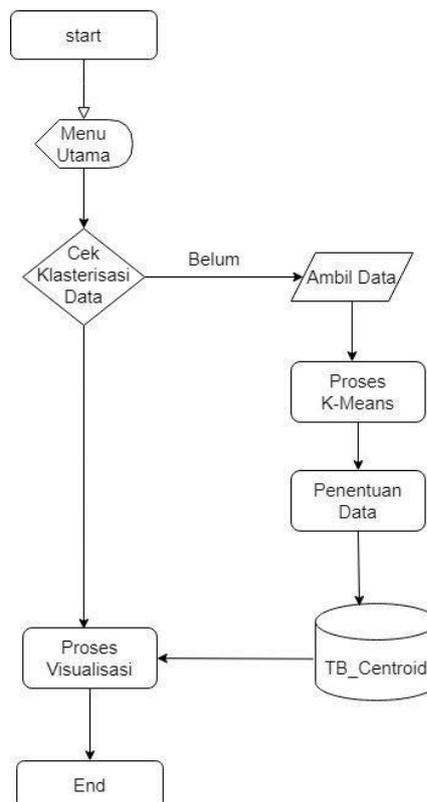
y_i : banyaknya data y ke-i

n : banyaknya data

- d. Mengelompokkan data cluster, berdasarkan jarak terdekat pada centroid.
- e. Mengulangi cara 3-4, lakukan hingga data tidak ada yang berpindah pada cluster lain.

2.2. Desain Sistem (Arsitektur)

2.2.1. Flowchart



Gambar 1. Flowchart Sistem

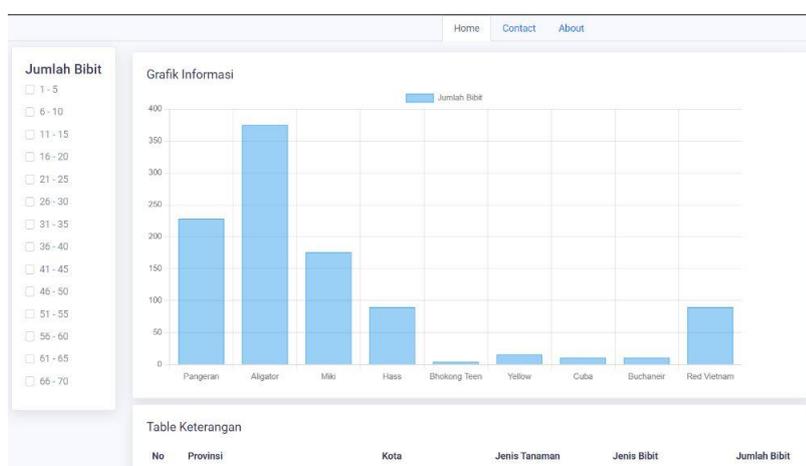
Flowchart merupakan bagan-bagan yang menggambarkan langkah-langkah atau proses dalam penyelesaian suatu masalah[10]. Pada gambar 1 tersebut menjelaskan flowchart sistem visualisasi data dimana pemilik akan login ke halaman utama kemudian memilih data yang sudah

di *cluster* menggunakan metode *K-Means* untuk mencari nilai *SSE*. Jika data belum di *cluster* maka pemilik akan mengambil data kemudian mengolahnya menggunakan metode *K-Means*. Kemudian data yang sudah ditentukan akan ditampung dalam tabel *centroid*. Selanjutnya akan dilakukan proses visualisasi data akan tetapi apabila data dari awal sudah di *cluster* maka pemilik akan langsung melakukan proses visualisasi data.

3. HASIL DAN PEMBAHASAN

3.1. Implementasi

Aplikasi yang telah peneliti buat merupakan hasil implementasi visualisasi data yang nantinya digunakan untuk melihat informasi mengenai pembeli bibit alpokat pada Pesantren Al-Fukaat.



Gambar 2. Tampilan Menu Utama Aplikasi

Pada gambar 2 terdapat desain menu aplikasi dimana desain *home* terdapat penggambaran grafik mengenai hasil data segmentasi pembeli berdasarkan data transaksi penjualan bibit alpokat. Pada menu lainnya yaitu *contact* didalamnya terdapat nomor dan *email* penjual yang dapat dihubungi langsung oleh pembeli sedangkan pada menu *about* adalah menu yang berisi tentang Pesantren Al-Fukaat.

3.2. Analisa Data

Pada penelitian data yang diolah sebagai adalah data pembelian bibit dan data pembeli bibit alpokat pada bulan Januari sampai bulan Mei 2022 yang bersumber pada Pesantren Al-Fukaat. Sampel data pembelian bibit alpokat dapat dilihat pada Tabel 1 dan Tabel 2 merupakan sampel data pembeli bibit alpokat.

Tabel 1. Sampel Data Pembelian Bibit Pada Bulan Januari 2022

No	Jenis Bibit	Jumlah
1	Pangeran	2
2	Aligator	6
3	Miki	5
4	Hass	2
5	Bhokhong Teen	1
6	Yellow	1
7	Cuba	1
8	Buchaneir	1
9	Red Vietnam	1
10	Markus	3

Tabel 2. Sampel Data Pembeli Bibit Alpokat Pada Bulan Januari 2022

No	Kota	Jumlah Pembeli	Jenis Tanah	Lingkungan
1	Semarang	2	Aluvial	Pedesaan
2	Kediri	20	Aluvial	Pedesaan dan perkotaan
3	Nganjuk	22	Aluvial	Pedesaan
4	Jombang	9	Aluvial	Pedesaan dan perkotaan
5	Mojokerto	2	Aluvial	Perkotaan
6	Sidoarjo	7	Aluvial	Perkotaan

3.3. Simulasi Algoritma

Data yang diolah adalah data pembeli dan pembelian bibit alpokat pada bulan Januari 2022 seperti pada Tabel 1 dan 2 kemudian dikelompokkan menggunakan metode *K-Means Clustering*.

a) Menentukan *cluster*

Langkah pertama adalah menentukan nilai *cluster* (*k*) yang diperoleh dari hasil data pembelian dengan jumlah 10 jenis bibit, dengan perhitungan *sturges* dapat dilihat pada persamaan (3) :

$$K = 1 + 3,3 * \text{Log} (N)$$

$$K = 1 + 3,3 * \text{Log} (10)$$

$$K = 4,3 \text{ dibulatkan menjadi } 4 \dots \dots \dots (3)$$

Untuk *cluster* pembeli terdapat 24 kota dengan perhitungan *struges* dapat dilihat pada persamaan (4) :

$$K = 1 + 3,3 * \text{Log} (N)$$

$$K = 1 + 3,3 * \text{Log} (24)$$

$$K = 5,9349 \text{ dibulatkan menjadi } 6 \dots \dots \dots (4)$$

Jadi jumlah *cluster* ada 4 *cluster* yang terdapat pada *cluster* pembelian dan terdapat 6 *cluster* pada *cluster* pembeli.

b) Menentukan *centroid*

Setelah diketahui banyaknya *cluster* maka dilanjutkan untuk menentukan *centroid* secara acak.

Tabel 3. *Centroid* awal data pembelian bibit

Jenis Bibit	Tanggal									
	1	2	3	4	5	6	7	8	...	31
Pangeran	1	2	2	2	2	1	1	1	...	2
Hass	2	0	0	0	2	1	1	0	...	0
Cuba	2	1	0	1	1	1	1	0	...	1
Markus	1	1	0	1	0	0	1	0	...	0

Tabel 4. *Centroid* awal data pembeli bibit alpokat

Kota	Jumlah Pembeli	Jenis Tanah	Lingkungan
Kediri	1	2	1
Tulungagung	1	1	2
Pasuruan	3	10	1
Ngawi	9	4	3
Kalurang	3	7	1
Banjarnegara	20	4	3

- c) Menghitung objek ke titik pusat (*centroid*)
- f. Menghitung *centroid* dapat dilihat pada persamaan (5) :

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3 \dots n \dots \dots \dots (5)$$

X dimaksudkan pada data bulan Januari 2022 dengan mengolah data pembelian dan pembeli bibit alpokat dengan *centroid cluster* yang ditentukan secara acak.

- d) Pada langkah perhitungan *cluster* menggunakan *centroid* awal pada data pembelian bibit alpokat hasil yang tidak berubah didapatkan dari hasil iterasi ke-3 dan iterasi ke-4 pada data pembeli bibit alpokat.

Tabel 5. Hasil akhir *cluster* data pembelian bibit

Jenis Bibit	C1	C2	C3	C4
Pangeran			✓	
Aligator	✓			
Miki	✓			
Hass		✓		
Bhokong Teen				✓
Yellow				✓
Cuba				✓
Buchaneir				✓
Red Vietnam	✓			
Markus	✓			

Tabel 6. Hasil akhir *cluster* data pembelian bibit

Kota	C1	C2	C3	C4	C5	C6
Semarang	✓					
Kediri						✓
Nganjuk						✓
Jombang				✓		
Mojokerto	✓					
Sidoarjo				✓		

- e) Langkah selanjutnya adalah menentukan hasil *SSE* menggunakan persamaan (6) :

$$\sum_{k=1}^k \sum_{x_i \in S_k} \|X_1 - C_k\|_2^2 \dots \dots \dots (6)$$

K dimaksudkan dengan jumlah *cluster* data pembelian dan data pembeli pada bulan Januari 2022 yang sudah ada pada tabel hasil akhir *cluster*.

Tabel 7. Hasil SSE data pembelian bibit alpokat

Jenis Bibit	m1	m2	m3	m4	Cluster	SEE
Pangeran	9,33	3,46	0	7,16	3	0
Hass	7,52	16,67	15,93	19,15	1	56,62
Cuba	4,91	13,45	12,68	15,88	1	24,12
Markus	10,05	7,68	3,46	6,73	2	0

Tabel 8. Hasil SSE data pembeli bibit alpokat

Kota	m1	m2	m3	m4	m5	m6	Cluster	SEE
Semarang	0,51	3,10	5,5	4,63	2,16	19,02	1	0,26
Kediri	18,33	19,06	18,92	13,64	16,93	1,41	6	2
Nganjuk	20,25	20,96	20,74	15,66	18,83	1,41	6	2
Jombang	7,29	8,49	9,12	2,83	6,25	12,04	4	8,04
Mojokerto	0,82	3,18	5,59	4,47	2,27	19	1	0,68
Sidoarjo	5,25	6,52	7,5	1,01	4,20	14	4	1,04

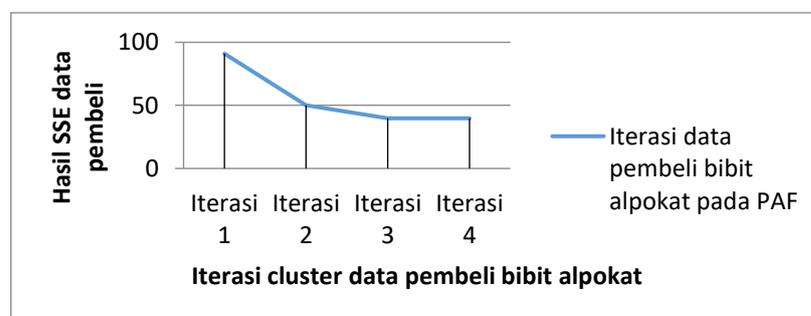
- f) Hasil SSE pembelian bibit alpokat diperoleh pada data iterasi ke-3 dan hasil iterasi ke-4 pada data pembeli bibit alpokat. Penyajian hasil dibawah ini berupa tabel dan grafik. Hasil visualisasi dari hasil perhitungan nilai SSE pada data pembelian bibit alpokat dan data pembeli bibit alpokat pada Pesantren Al-Fukaat bulan Januari 2022 dapat dilihat pada gambar 3 dan gambar 4.

Tabel 9. Hasil perhitungan SEE pembelian bibit alpokat pada Pesantren Al-Fukaat

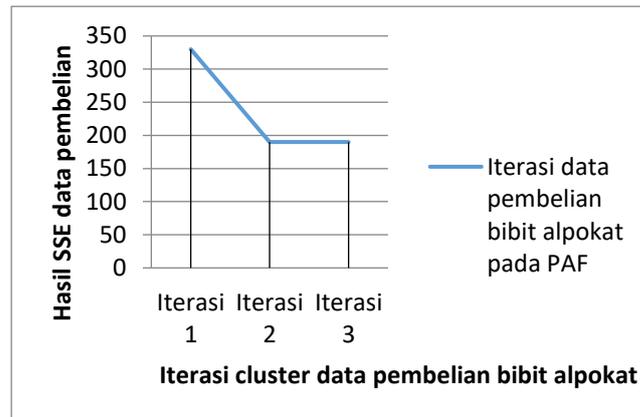
Iterasi	Hasil SSE
1	330
2	190
3	190

Tabel 10. Hasil perhitungan SEE pembeli bibit alpokat pada Pesantren Al-Fukaat

Iterasi ke-1	Hasil SSE
1	91
2	50,06
3	39,78
4	39,78



Gambar 3. Tampilan grafik hasil nilai SSE data pembeli bibit alpokat pada Pesantren Al-Fukaat



Gambar 4. Tampilan grafik hasil nilai *SSE* data pembelian bibit alpokate pada Pesantren Al-Fukaat

Berdasarkan hasil *SEE* yang diperoleh pada data pembelian bibit alpokate, iterasi *cluster* ke-3 lebih optimal dibanding iterasi ke-1 dan iterasi ke-2 yaitu dengan hasil *SSE* 190. Sedangkan pada data pembeli bibit alpokate hasil iterasi ke-4 hasilnya lebih rendah dibandingkan iterasi ke-1, iterasi ke-2 dan iterasi ke-3, yaitu dengan hasil *SSE* 39,7852. Oleh sebab itu aplikasi segmentasi pembeli menggunakan metode *K-Means* dapat digunakan karena dapat menghasilkan nilai *SSE* yang lebih rendah.

4. KESIMPULAN

4.1. Simpulan

Hasil dari penelitian segmentasi pembeli bibit alpokate menggunakan metode *K-Means Clustering* untuk pengelompokan data pembeli pada Pesantren Al-Fukaat menghasilkan pengelompokan pembeli berdasarkan pembelian bibit alpokate dan banyaknya pembelian bibit pada setiap daerah. Dari data pembelian didapatkan 10 jenis bibit yang diteliti dan didapat 4 *cluster*. *Cluster* 1 ada 4 jenis bibit, *cluster* 2 sebanyak 1 jenis bibit, *cluster* 3 sebanyak 1 jenis bibit, *cluster* 4 sebanyak 4 jenis bibit. Pada data pembeli didapatkan 24 kota dari pembeli terdapat 6 *cluster*. *Cluster* 1 ada 7 kota, *cluster* 2 ada 5 kota, *cluster* 3 ada 2 kota, *cluster* 4 sebanyak 5 kota, *cluster* 5 sebanyak 3 kota, *cluster* 6 ada 2 kota.

Berdasarkan hasil *SEE* yang diperoleh pada data pembelian bibit alpokate, iterasi *cluster* ke-3 lebih optimal dibanding iterasi ke-1 dan iterasi ke-2 yaitu dengan hasil *SSE* 190. Sedangkan pada data pembeli bibit alpokate hasil iterasi ke-4 hasilnya lebih optimal dibandingkan iterasi ke-1, iterasi ke-2 dan iterasi ke-3, yaitu dengan hasil *SSE* 39,7852. Oleh sebab itu aplikasi segmentasi pembeli menggunakan metode *K-Means* dapat digunakan karena dapat menghasilkan nilai *SSE* yang lebih rendah. Hal tersebut mempermudah penjual dalam menentukan target pembeli bibit alpokate.

4.2. Saran

Berdasarkan penelitian yang telah dilakukan, maka saran sangat diharapkan penulis dalam pengembangan selanjutnya mengenai aplikasi segmentasi pembeli dengan berbagai metode lainnya.

DAFTAR PUSTAKA

- [1] P. Studi, T. Industri, and S. Vokasi, "PRODUKSI DAN PEMASARAN BENIH ALPUKAT (*Persea americana*) VARIETAS WINA DI PT BOTANI SEED INDONESIA," 2021.

- [2] P. H. Suharti, A. S. Suryandari, and R. N. Amalia, “Analisis Kinerja Modul Pengendali Tekanan Udara Pct-14 Berbasis Plc Dengan Berbagai Metoda Tuning,” *Sebatik*, vol. 26, no. 2, pp. 420–427, 2022, doi: 10.46984/sebatik.v26i2.2134.
- [3] A. Sumeke, S. L. H V Joyce Lopian, and S. S. . Loindong, “Analisis Segmentasi Demografi Dan Psikografi Pada Toko Raps,” *J. EMBA*, vol. 10, no. 1, pp. 202–207, 2022, [Online]. Available: <https://ejournal.unsrat.ac.id/index.php/emba/article/view/37680>
- [4] Suryadi Muzahidi Aziz and Nur Azizah Komara Rifai, “Pengelompokkan Ekspor Kopi Menurut Negara Tujuan Menggunakan Metode K-Means Clustering dengan Silhouette Coefficient,” *Bandung Conf. Ser. Stat.*, vol. 2, no. 2, pp. 416–424, 2022, doi: 10.29313/bcss.v2i2.4536.
- [5] A. Pradana, A. Sopiyan, T. A. Pratama, and A. Yusuf, “Analisis strategi pemasaran gopay berupa segmentasi, targeting, dan positioning serta pengaruhnya terhadap keputusan pembelian,” *Online) J. Manaj.*, vol. 14, no. 3, pp. 2022–551, 2022, doi: 10.29264/jmmn.v14i3.11443.
- [6] A. Karami and Y. Widharto, “PERANCANGAN BUSINESS INTELLIGENCE DAN SEGMENTASI PELANGGAN MENGGUNAKAN K MEANS CLUSTERING BERDASARKAN RFM MODEL”.
- [7] N. Mirantika, T. S. Syamfithriani, and R. Trisudarmo, “Implementasi Algoritma K-Medoids Clustering Untuk Menentukan Segmentasi Pelanggan,” *J. Nuansa Inform.*, vol. 17, pp. 2614–5405, 2023, [Online]. Available: <https://journal.uniku.ac.id/index.php/ilkom>
- [8] S. O. Hayati, “Analisis Segmentasi Pasar Terhadap Volume Penjualan Melalui Keputusan Pembelian:(Survei Pada Toko Youth Sneakers),” *J-CEKI J. Cendekia Ilm.*, vol. 1, no. 5, pp. 614–622, 2022, [Online]. Available: <http://ulilalbabinstitute.com/index.php/J-CEKI/article/view/761>
- [9] J. Maraja *et al.*, “Penerapan Metode Clustering Dalam Menentukan Persentase Penyakit Pada Pasien Dengan Menggunakan Algoritma K-Means Pada Puskesmas,” no. x, 2020.
- [10] V.A.R.Barao, R.C.Coata, J.A.Shibli, M.Bertolini, and J.G.S.Souza, “No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title,” *Braz Dent J.*, vol. 33, no. 1, pp. 1–12, 2022.

