

PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA PENJUALAN VOUCHER INTERNET STUDY KASUS KONTER ABC

¹Mohammad Sabar Jamil, ²Nugraha Yudhasyah, ³Hilman Mutaqin, ⁴Moh. Milki I. M.

^{1,2,4}Program Studi Informatika Sekolah Tinggi Teknologi Cipasung

³Program Studi Sistem Informasi STMIK LIKMI Bandung

Jl. Raya Singaparna-Ciawi Km.1 PO. BOX 24 Cilampunghilir Padakembang Tasikmalaya 46466

Telp./Fax : 0265-2550424

Email: jamil@sttcipasung.ac.id

nugrahayudhasyah@sttcipasung.ac.id

10219005@sttcipasung.ac.id

Abstract— *Rapid technological advances will bring changes to aspects of human life, one of which is how to communicate. With the discovery of cell phones, it is easier for humans to communicate remotely, one of which is WhatsApp. If someone wants to use WhatsApp, internet access is needed so that the sales of internet vouchers are increasing. Data mining is the process of processing data in order to obtain new information, clustering is chosen because it aims to create clusters from existing data. This study aims to cluster internet vouchers so that counter owners can make stock vouchers more precisely. The results of this study yielded the highest value for cluster C1 35,000 while the lowest value for C2 was 43,000 with a total of 99 data with a percentage of 95.19% while cluster C2 consisted of 5 data with a percentage of 4.81%. the conclusion is that the most purchased nominal vouchers are under 40,000. therefore the counter owner can keep more stock for a nominal value below 40,000*

Keywords— *Clustering, K-Means, Voucher Internet*

Abstrak— Kemajuan teknologi yang pesat akan membawa perubahan pada aspek kehidupan manusia yang salah satunya adalah cara berkomunikasi. Dengan ditemukannya telepon seluler, maka memudahkan manusia untuk melakukan komunikasi jarak jauh salah satunya whatsapp. Jika seseorang ingin menggunakan whatsapp, maka diperlukan akses internet sehingga penjualan voucher internet semakin bertambah banyak. Data mining adalah proses pengilahan data agar mendapatkan suatu informasi baru, clustering dipilih karena bertujuan untuk membuat cluster dari data yang ada. Penelitian ini bertujuan melakukan clustering voucher internet sehingga pemilik konter bisa membuat stok voucher dengan lebih tepat. Hasil penelitian ini menghasilkan nilai tertinggi cluster C1 35.000 sedangkan nilai terendah C2 43.000 dengan jumlah data 99 data dengan persentase 95.19% sedangkan cluster C2 berjumlah 5 data dengan persentase 4.81%. kesimpulannya adalah bahwa nominal voucher yang paling banyak dibeli dibawah 40.000. oleh karena itu pemilik konter bisa menyimpan stok lebih banyak untuk nominal dibawah 40.000

Kata kunci : Clustering, K-Means, Voucher Internet

I. PENDAHULUAN

Kemajuan teknologi yang sangat pesat akan membawa perubahan pada semua aspek kehidupan manusia yang sebelumnya kurang modern menjadi modern [1]. Salah satu yang berubah adalah cara manusia berkomunikasi, jika dahulu hanya menggunakan telepon umum jika manusia

ingin saling berkomunikasi jarak jauh, namun sekarang dengan adanya telepon seluler dimana dan kapan saja bisa berkomunikasi. Karena penggunaan telepon seluler menjadi semakin sering digunakan mana jumlah telepon seluler semakin banyak.

Penggunaan telepon seluler yang semakin banyak menyebabkan data transaksi antara

pembeli dan konter semakin banyak. Konter akan mempunyai data transaksi yang sangat banyak karena semakin banyak orang yang menggunakan telepon seluler untuk berkomunikasi khususnya Whatsapp. Untuk dapat menggunakan whatsapp, diperlukan akses internet sehingga penjualan voucher internet semakin bertambah banyak. Transaksi adalah pertukaran barang ataupun jasa yang dilakukan oleh seseorang ataupun suatu organisasi [2].

Konter ABC merupakan salah satu konter yang berlokasi di Kota Tasikmalaya yang menjual berbagai saldo seperti saldo listrik, saldo angkutan *online*, uang digital, voucher internet, dan pulsa. Hampir setiap hari konter ABC melakukan ratusan transaksi sehingga jumlah data yang transaksi yang dibuat bisa mencapai ribuan perbulan yang direkap pada file Microsoft Excel. Namun ribuan data tersebut tidak diolah menjadi sehingga data-data tersebut menjadi tidak bermanfaat karena tidak memberikan informasi apapun. Data yang hanya tertumpuk dan tidak diolah tidak akan menghasilkan suatu informasi sehingga tidak memberikan manfaat [3]. Namun jika data data yang tertumpuk itu diproses dan diolah maka akan menghasilkan suatu informasi [4].

Tujuan penelitian ini adalah melakukan klusterisasi penjualan voucher internet berdasarkan moninal agar diperoleh suatu informasi yang berguna untuk pemilik konter sehingga diharapkan dapat membantu dalam pengambilan keputusan seperti membuat kebijakan persediaan.

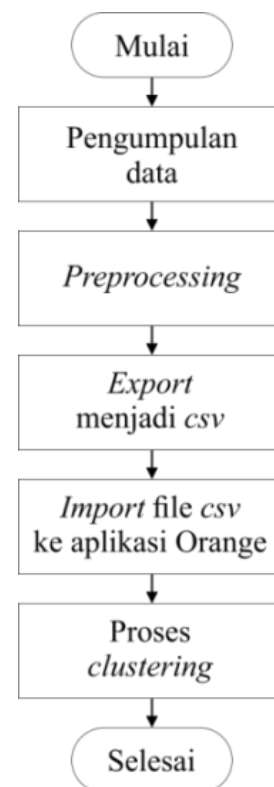
II. LANDASAN TEORI

Data mining merupakan suatu proses untuk menemukan suatu pola yang menarik serta pengetahuan dari suatu data yang besar. Sumber data bisa berupa gudang data, basis data, maupun sistem dinamis [5]. Juga dipakai untuk mendapatkan suatu informasi yang baru serta berguna dari jumlah data yang cukup besar [6]. *Clustering* adalah salah satu teknik dari data mining yang bertujuan untuk mengelompokan suatu data berdasarkan suatu kesamaan-kesamaan yang ada pada suatu data [7]. *K-Means* merupakan suatu algoritma yang cukup mudah serta efektif untuk menemukan cluster yang ada dalam data [8]. Orange adalah suatu aplikasi *open source* yang

dapat digunakan untuk melakukan analisis serta visualisasi data. Aplikasi ini disukai karena faktor kendalan serta inovatif. Aplikasi ini juga mempermudah penggunaannya karena menyajikan informasi dalam bentuk visual serta intuitif [9].

III. METODE PENELITIAN

Penelitian dilakukan dengan menggunakan data sekunder serta jenisnya adalah kuantitatif. *Flowchart* penelitian bisa dilihat pada gambar dibawah ini.



Gambar 1. *Flowchart* penelitian

IV. HASIL PENELITIAN

1. Pengumpulan data

Data yang didapat adalah data penjualan voucher internet dari konter ABC pada tanggal 2 Agustus tahun 2022 sebanyak 114 data. Data diperoleh dari wawancara dengan pemilik konter serta mengumpulkan dari transaksi yang berlangsung pada tanggal tersebut. Namun data yang didapat tidak bisa langsung diolah karena ada beberapa data yang tidak termasuk kedalam transaksi voucher

internet. Tabel 1 merupakan data penjualan konter ABC

Tabel I
Dataset penjualan

Tanggal	Jenis	Harga
02 Agustus 2022	sp tsel 35	85,000
	v tsel 10	65,000
	v tsel 1.5	8,000
	v tsel 1.5	8,000
	v tsel 1.5	8,000
	v tsel 1.5	8,000
	v tsel 1.5	8,000
	v tsel 1.5	8,000

	kabel robot	15,000
	kabel robot	15,000
	kabel robot	15,000
	headset jbl	15,000

Tabel II
Dataset yang telah dibersihkan

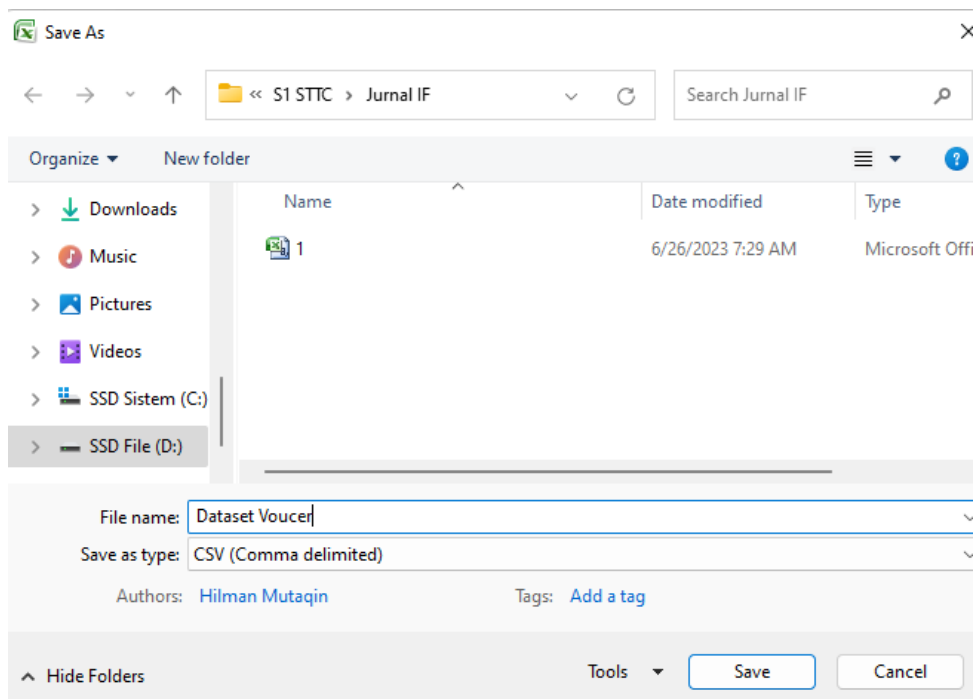
No	Harga
1	85000
2	17000
3	17000
4	10000
5	10000
6	15000
7	20000
8	20000
9	52000
10	15000
11	15000
12	15000
13	35000
14	35000
15	35000
16	7000
17	7000
18	7000
19	13000
20	13000
.....
.....
.....
104	20000

2. Preprocessing

Dataset yang sudah diperoleh tidak bisa langsung diproses, melainkan harus terlebih dahulu dibersihkan seperti menghapus kolom tanggal dan menghapus transaksi selain penjualan voucher internet. Setelah data dibersihkan hanya ada 104 data penjualan voucher internet. Berikut adalah dataset penjualan voucher internet yang sudah dibersihkan.

3. Export menjadi file CSV

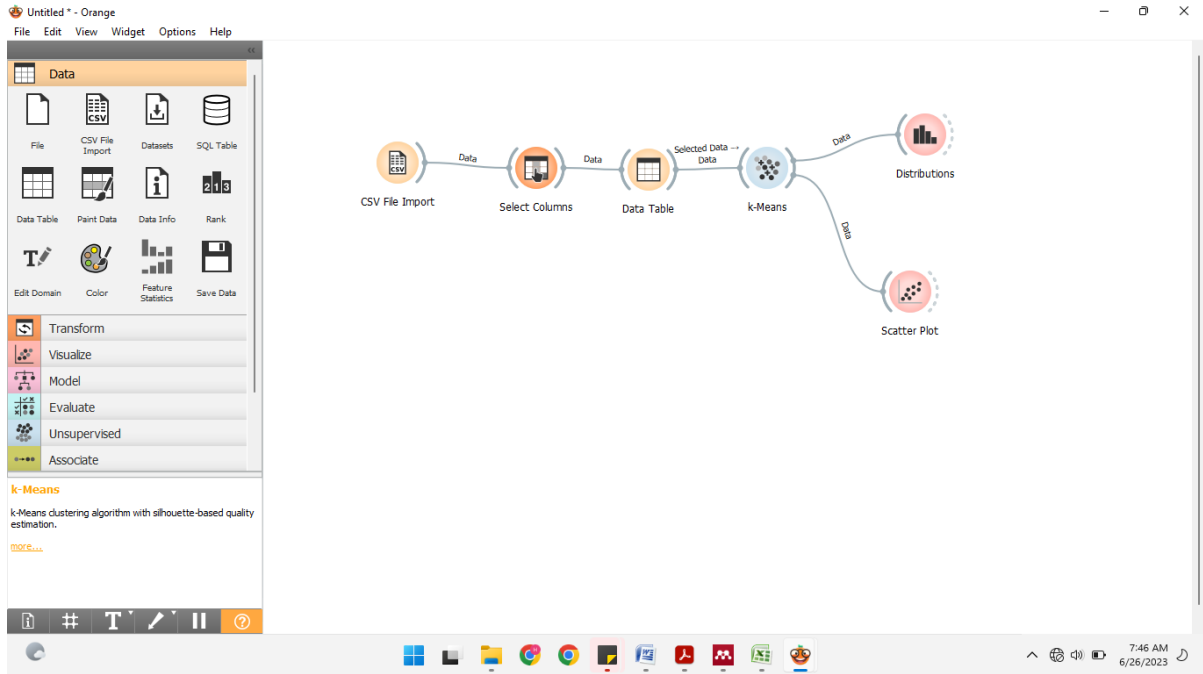
Setelah data dibersihkan, *export* menjadi file CSV



Gambar 1. *Export file CSV*

Buat format alat seperti dibawah ini

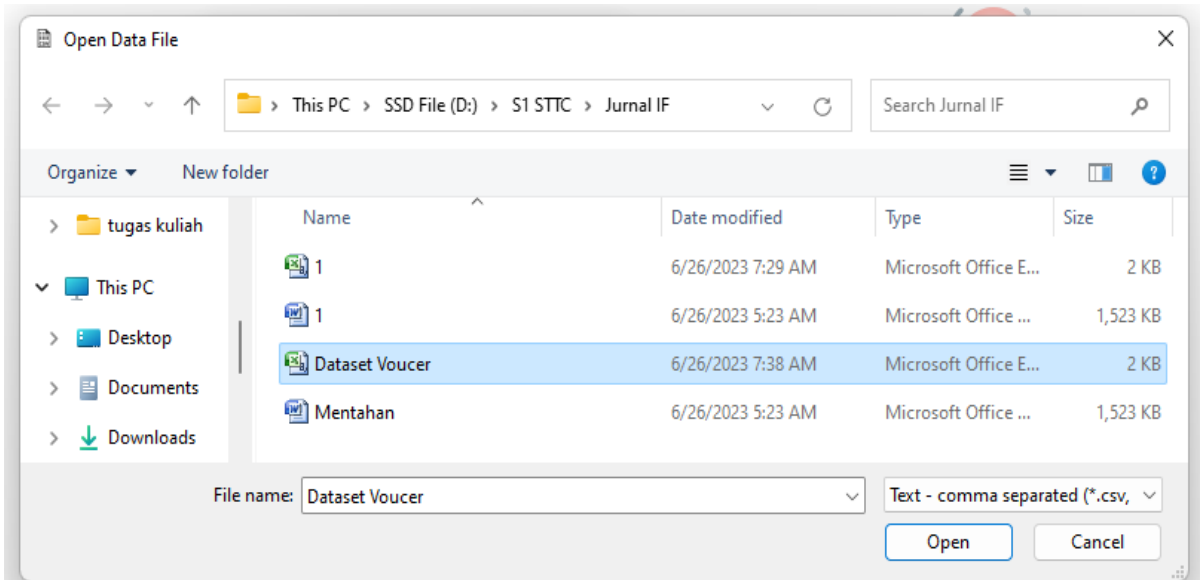
4. Persiapan alat



Gambar 2. Persiapan alat

5. *Import file SCV* kedalam aplikasi Orange

Setelah *diexport*, selanjutnya *import* kedalam aplikasi orange

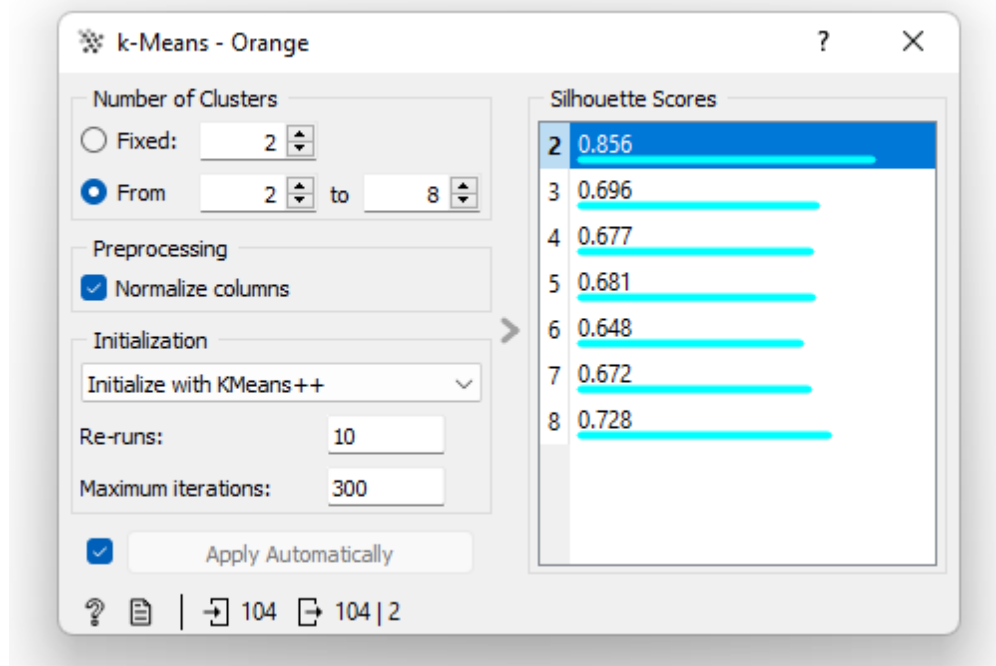


Gambar 3. *Import file* kedalam aplikasi Orange

6. Proses *clustering*

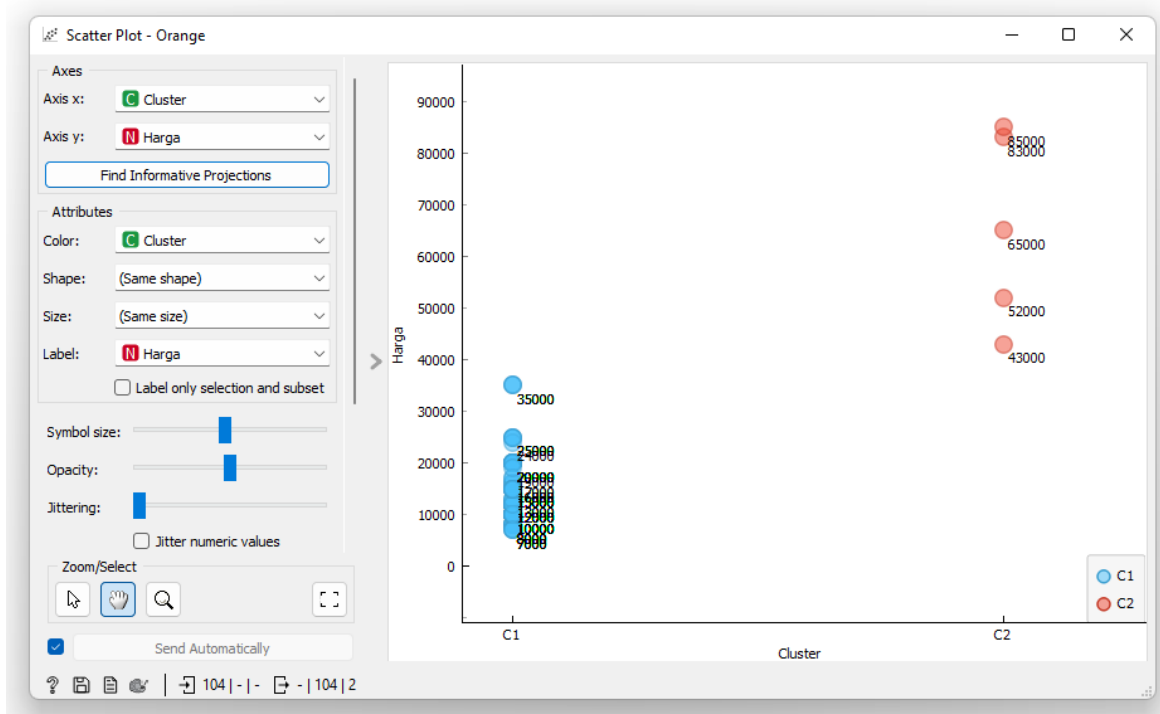
Clustering dilakukan otomatis oleh aplikasi. Dari penelitian yang

dilakukan, jumlah cluster yang dibuat sebanyak 2 dengan nilai silhouette 0.856



Gambar 4. Pemilihan jumlah *cluster*

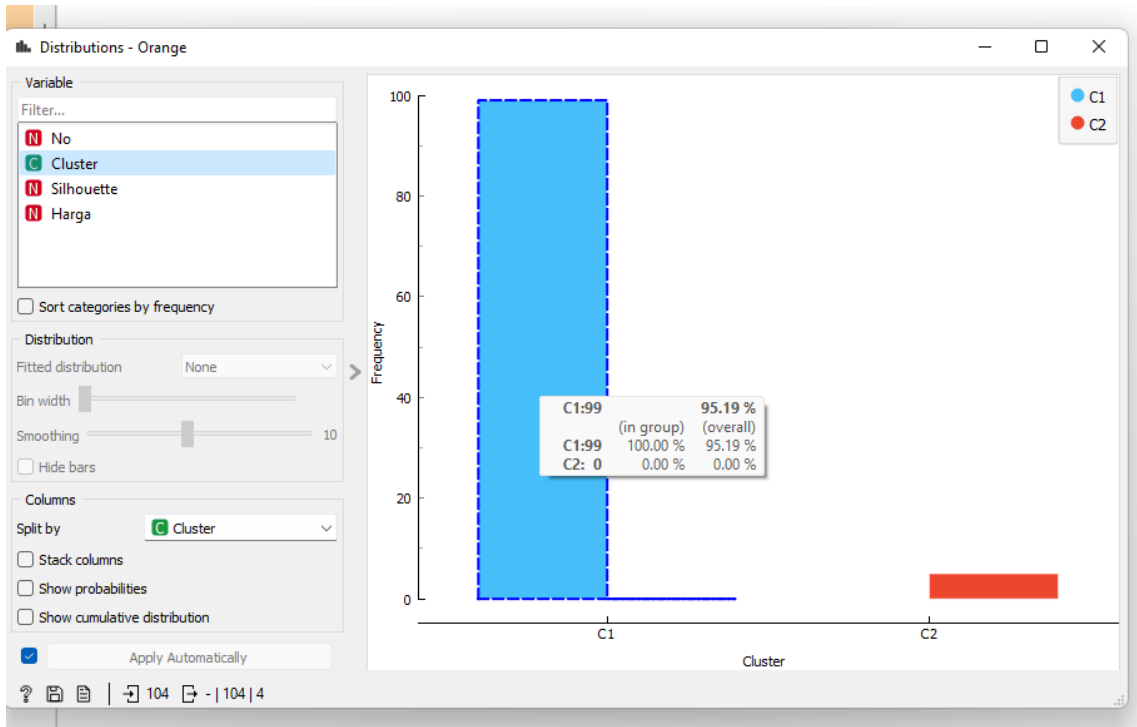
Hasil penyebaran cluster dengan scatter plot dapat dilihat dari gambar berikut ini



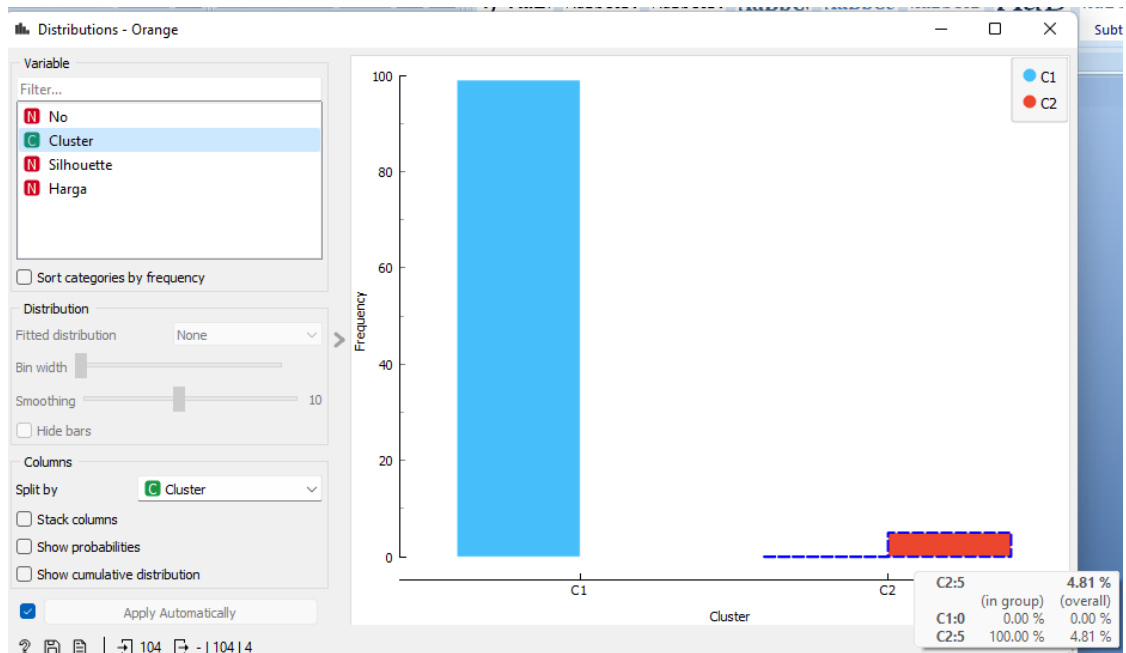
Gambar 5. Scatter plot

Dari gambar diatas dapat dilihat bahwa nilai tertinggi cluster C1 35.000 sedangkan nilai terendah C2 43.000. jika dilihat dari jumlah data, cluster C1 berjumlah 99 data

dengan persentase 95.19% sedangkan cluster C2 berjumlah 5 data dengan persentase 4.81%



Gambar 5. Jumlah cluster C1



Gambar 6. Jumlah cluster C2

V. KESIMPULAN

Dari hasil pengolahan data yang telah dilakukan, dapat disimpulkan bahwa nominal voucher yang paling banyak dibeli dibawah 40.000. oleh karena itu pemilik konter bisa

menyimpan stok lebih banyak untuk nominal dibawah 40.000

DAFTAR PUSTAKA

- [1] N. Huda, "Analisis Kinerja Website Pt Pln (Persero) Menggunakan Metode Pieces," *Sistemasi*, vol. 8, no. 1, pp. 78–

- 89, 2019, doi: DOI: <https://doi.org/10.32520/stmsi.v8i1>.
- [2] P. Fakhriyah, "Pengaruh Layanan Transportasi Online (Gojek) Terhadap Perluasan Lapangan Kerja Bagi Masyarakat Di Kota Cimahi," *Comm-Edu (Community Educ. Journal)*, vol. 3, no. 1, p. 34, 2020, doi: 10.22460/comm-edu.v3i1.3719.
- [3] Nawassyarif, M. Julkarnain, and K. Rizki Ananda, "Sistem Informasi Pengolahan Data Ternak Unit Pelaksana Teknis Produksi Dan Kesehatan Hewan Berbasis Web," *J. Inform. Teknol. dan Sains*, vol. 2, no. 1, pp. 32–39, 2020, doi: 10.51401/jinteks.v2i1.556.
- [4] A. Syafnur, "Analisis Dengan Metode Klasifikasi Menggunakan Decision Tree Untuk Memprediksi Penentuan Resiko kredit Bank," *Jurteks*, vol. 4, no. 1, pp. 101–106, 2017, doi: 10.33330/jurteks.v4i1.30.
- [5] N. Agustina and P. Prihandoko, "Perbandingan Algoritma K-Means dengan Fuzzy C-Means Untuk Clustering Tingkat Kedisiplinan Kinerja Karyawan," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 3, pp. 621–626, 2018, doi: 10.29207/resti.v2i3.492.
- [6] Yoga Religia, Agung Nugroho, and Wahyu Hadikristanto, "Klasifikasi Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 187–192, 2021, doi: 10.29207/resti.v5i1.2813.
- [7] T. Syahputra, J. Halim, and E. P. Sintho, "Penerapan Data Mining Dalam Menentukan Pilihan Jurusan Bidang Studi SMA Menggunakan Metode," *Penerapan Data Min. dalam Menentukan Pilihan Jur. di Bid. Stud. SMA menggunakan Metod. Clust. Dengan Tek. Single Link. JURTEKSI*, vol. IV, no. 2, pp. 1–4, 2018.
- [8] G. A. Marcoulides, *Discovering Knowledge in Data: an Introduction to Data Mining*, vol. 100, no. 472. 2005. doi: 10.1198/jasa.2005.s61.
- [9] R. A. raffaidy Wiguna and A. I. Rifai, "Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining," *J. Inf. Syst. Informatics*, vol. 3, no. 1, pp. 1–12, 2021, doi: 10.33557/journalisi.v3i1.78.