

**ANALISIS DATA PENJUALAN PADA TOKO BUKU SOCIAL
AGENCY BARU AMBARUKMO MENGGUNAKAN METODE
*K-MEANS CLUSTERING***

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan Mencapai Sarjana
S-1 Program Studi Teknik Informatika



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
Disusun oleh:
Mohammad Ikhwanud Dawam
19106050050

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA

2023

HALAMAN PENGESAHAN



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2316/Un.02/DST/PP.00.9/08/2023

Tugas Akhir dengan judul : ANALISIS DATA PENJUALAN PADA TOKO BUKU SOCIAL AGENCY BARU
AMBARUKMO MENGGUNAKAN METODE K-MEANS CLUSTERING

yang dipersiapkan dan disusun oleh:

Nama : MOHAMMAD IKHWANUD DAWAM
Nomor Induk Mahasiswa : 19106050050
Telah diujikan pada : Senin, 21 Agustus 2023
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Ir. Maria Ulfah Siregar, S.Kom., MIT., Ph.D.
SIGNED

Valid ID: 64e6fe950cff7



Penguji I

Nurochman, S.Kom., M.Kom
SIGNED

Valid ID: 64e6c0798b595



Penguji II

Muhammad Galih Wonoseto, M.T.
SIGNED

Valid ID: 64e6c02e647ec



Yogyakarta, 21 Agustus 2023
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 64e70ebd4199f

SURAT PERSETUJUAN SKRIPSI

Hal : Persetujuan Tugas Akhir

Lamp : -

Kepada:

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

Di Yogyakarta

Assalamu'alaikum Wr. Wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka saya selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Mohammad Ikhwanud Dawam
NIM : 19106050050
Judul Skripsi : Analisis Data Penjualan Pada Toko Buku Social Agency Baru Ambarukmo Menggunakan Metode *K-Means Clustering*

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika.

Dengan ini saya berharap agar tugas akhir Saudara dapat segera di-*munaqasyah*-kan. Atas perhatiannya kami ucapkan terimakasih.

Wassalamu'alaikum Wr. Wb.

Yogyakarta, 11 Agustus 2023
Pembimbing,



Ir. Maria Ulfah Siregar, S.Kom., MIT., Ph.D.
NIP. 19780106 200212 2 001

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini:

Nama : Mohammad Ikhwanud Dawam

NIM : 19106050050

Program Studi : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul “**Analisis Data Penjualan Pada Toko Buku Social Agency Baru Ambarukmo Menggunakan Metode *K-Means Clustering***” merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar Pustaka.

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAYA
YOGYAKARTA

Yogyakarta, 11 Agustus 2023



Mohammad Ikhwanud Dawam
NIM. 19106050050

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT, atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan judul "**Analisis Data Penjualan Pada Toko Buku Social Agency Baru Ambarukmo Menggunakan Metode *K-Means Clustering***". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi Universitas Islam Negeri Islam Sunan Kalijaga.

Dalam proses penulisan skripsi ini penulis ingin mengucapkan terima kasih yang sebesar-besarnya pihak-pihak yang telah membantu dalam penyelesaian skripsi baik langsung maupun tidak langsung. Oleh karena itu penulis ingin mengucapkan terima kasih sebesar-besarnya kepada:

1. Allah SWT yang telah memberikan saya karunia untuk menulis skripsi hingga selesai.
2. Nabi Muhammad SAW yang telah memberi saya tauladan dalam keseharian saya sehingga saya dapat menyelesaikan skripsi ini dengan baik.
3. Semua orang tua dan keluarga saya yang telah memberi saya doa, dukungan, dan semangat.
4. Ibu Dosen, Ir. Maria Ulfah Siregar, S. Kom., MIT., Ph.D., selaku Kepala Program Studi Teknik Informatika UIN Sunan Kalijaga dan selaku Pembimbing saya yang telah memberi bimbingan, arahan, dan masukan yang berharga dalam proses penulisan skripsi ini.
5. Bapak/Ibu Dosen Teknik Informatika UIN Sunan Kalijaga yang telah memberikan ilmu dan wawasan yang bermanfaat bagi saya dari semester awal hingga akhir.
6. Pihak Toko SAB Ambarukmo yang menerima saya dengan terbuka dan memberikan saya data yang sangat membantu dalam pengerjaan skripsi saya.
7. Teman-teman seperjuangan Teknik Informatika Angkatan 2019
8. Teman-teman Pondok Pesantren Minhajut Tamyiz

9. Jauharotillah yang telah memberikan penulis dukungan, semangat dalam mengerjakan skripsi ini, terima kasih atas kesabaran dalam menghadapi waktu-waktu yang penulis habiskan untuk menyelesaikan skripsi ini. Meskipun terkadang penulis sibuk dan terfokus pada penulisan dan hal lainnya, engkau tetap memberikan dukungan dan pengertian.
10. Semua pihak yang tidak dapat disebutkan satu per satu namun telah memberikan kontribusi positif dalam perjalanan penulisan skripsi ini.

Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik, saran, dan masukan yang membangun dari pembaca untuk perbaikan di masa yang akan datang. Semoga skripsi ini dapat memberikan kontribusi dan manfaat bagi pengembangan ilmu pengetahuan dan dunia informatika. Terima kasih.

Yogyakarta, Juni 2023

Penulis



Mohammad Ikhwanud Dawam

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Saya persembahkan penulisan skripsi saya kepada:

Allah SWT dan Rasulullah SAW

Yang telah memberi hidayah, menuntun, membimbing, dan memberi kekuatan saat suka maupun duka.

Orang tua saya, khususnya **Ayahanda Machfudz** dan **Ibu Farah Faida** yang selalu memberikan saya motivasi dan kasih sayang dalam mengerjakan skripsi

Dek El, Dek Aya selaku adikku tersayang

Jauharotillah yang selalu memberi motivasi dan semangat



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN MOTTO

“Allah menghendaki kemudahan bagimu dan tidak menghendaki kesukaran”

- QS. Al-Baqarah [2]:185

“Dengan bersabar hal-hal besar akan mudah tercapai”

- Ali bin Abi Thalib R.A

“Hidup bukanlah permainan keberuntungan. Jika kau ingin menang, kau harus bekerja keras.”

- Sora (No Game No Life)



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR ISI

HALAMAN PENGESAHAN.....	i
SURAT PERSETUJUAN SKRIPSI.....	ii
PERNYATAAN KEASLIAN SKRIPSI.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Penelitian.....	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
1.6 Keaslian Penelitian.....	4
1.7 Sistematika Penelitian.....	4
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	6
2.1 Tinjauan Pustaka.....	6
2.2 Landasan Teori.....	11
2.2.1 <i>Data Mining</i>	11
2.2.2 <i>Data Preprocessing</i>	13

2.2.3 Clustering.....	15
2.2.4 Algoritma <i>K-Means</i>	16
2.2.5 <i>K-Means++</i>	17
2.2.6 <i>Forgy Initialization</i>	19
2.2.7 <i>Euclidean Distance</i>	19
2.2.8 <i>Silhouette Coefficient</i>	19
2.2.9 <i>Python</i>	20
2.2.10 <i>Streamlit</i>	21
BAB III METODE PENELITIAN	22
3.1 Jenis dan Desain Penelitian.....	22
3.2 Lokasi dan Waktu Penelitian.....	22
3.3 Teknik Pengumpulan Data	22
3.4 Alat dan Bahan	23
3.5 Teknik Analisis Data	23
BAB IV HASIL DAN PEMBAHASAN	25
4.1 Hasil	25
4.1.1 <i>Gathering Data</i>	25
4.1.2 <i>Preprocessing Data</i>	26
4.1.2.1 <i>Normalisasi Data Mentah</i>	26
4.1.2.2 <i>Filtering Data</i>	29
4.1.2.3 <i>Sum Field Quantity</i>	30
4.1.2.4 <i>Membuat Field Total_Orders</i>	32
4.1.2.5 <i>Pembersihan Data Akhir</i>	33
4.1.3 <i>Clustering K-Means</i>	34
4.1.3.1 <i>Import Dataset</i>	34

4.1.3.2 Penentuan Jumlah <i>Cluster</i>	35
4.1.3.3 Penggunaan <i>K-Means++</i> dan <i>Forgy Initialization</i>	35
4.1.3.4 <i>Euclidean Distance</i>	39
4.1.3.5 <i>K-Means Clustering</i> dan Iterasi <i>Clustering</i>	40
4.2 Pembahasan.....	42
4.2.1 Hasil <i>Clustering</i>	42
4.2.2 <i>Silhouette Coefficient</i>	43
4.2.3 Labelisasi <i>Cluster</i> dan Visualisasi Hasil	44
4.2.3.1 Metode Inisialisasi <i>K-Means++</i> Dengan 3 <i>Cluster</i>	45
4.2.3.2 <i>Forgy Initialization</i> Dengan 3 <i>Cluster</i>	47
4.2.3.3 <i>K-Means++</i> Dengan 2 <i>Cluster</i>	50
4.2.3.4 <i>Forgy Initialization</i> Dengan 2 <i>Cluster</i>	52
4.2.4 Perbandingan Hasil <i>Clustering</i>	54
BAB V PENUTUP.....	55
5.1 Kesimpulan	55
5.2 Saran.....	55
DAFTAR PUSTAKA.....	57
CURRICULUM VITAE	60

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka.....	8
Tabel 4. 1 Metode <i>K-Means++</i> 3 Cluster Awal ([5, 5], [21, 523], [9, 159])	43
Tabel 4. 2 Metode <i>Forgy Initialization</i> 3 Cluster Awal ([2, 3], [10, 10], [2, 2]) ..	43
Tabel 4. 3 Metode <i>K-Means++</i> 2 Cluster Awal ([5, 5], [21, 523])	43
Tabel 4. 4 Metode <i>Forgy Initialization</i> 2 Cluster Awal ([3, 3], [2, 2])	43
Tabel 4. 5 Tabel Perhitungan <i>Silhouette Score</i>	44
Tabel 4. 6 <i>K-Means++</i> (3 Cluster) Klaster 1 (Kurang Laris)	46
Tabel 4. 7 <i>K-Means++</i> (3 Cluster) Klaster 2 (Laris)	47
Tabel 4. 8 <i>K-Means++</i> (3 Cluster) Klaster 3 (Cukup Laris).....	47
Tabel 4. 9 <i>Forgy Initialization</i> (3 Cluster) Klaster 1 (Cukup Laris)	48
Tabel 4. 10 <i>Forgy Initialization</i> (3 Cluster) Klaster 2 (Laris)	49
Tabel 4. 11 <i>Forgy Initialization</i> (3 Cluster) Klaster 3 (Kurang Laris)	49
Tabel 4. 12 <i>K-Means++</i> (2 Cluster) Klaster 1 (Kurang Laris)	51
Tabel 4. 13 <i>K-Means++</i> (2 Cluster) Klaster 2 (Laris)	51
Tabel 4. 14 <i>Forgy Initialization</i> (2 Cluster) Klaster 1 (Laris)	52
Tabel 4. 15 <i>Forgy Initialization</i> (2 Cluster) Klaster 2 (Kurang Laris)	53
Tabel 4. 16 Perbandingan Hasil <i>Clustering</i>	54

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

Gambar 2. 1 Langkah-Langkah <i>Data Mining</i> Dalam <i>Knowledge Discovery</i> (Han et al., 2012).....	12
Gambar 2. 2 Bentuk Dari <i>Preprocessing Data</i> (Han et al., 2012).....	15
Gambar 2. 3 Hasil <i>Clustering</i> Kurang Optimal.....	17
Gambar 2. 4 Hasil <i>Clustering</i> Lebih Optimal.....	18
Gambar 3. 1 Teknik Analisis Data.....	24
Gambar 4. 1 Tampilan Sampel Data Penjualan Toko SAB Ambarukmo	25
Gambar 4. 2 Data Sebelum Normalisasi	27
Gambar 4. 3 Data Setelah Normalisasi.....	27
Gambar 4. 4 Perbedaan No Faktur dan Kode Barang	28
Gambar 4. 5 Data Penjualan Berformat CSV Dari Data Excell.....	29
Gambar 4. 6 <i>Source Code</i> Proses Filtering Data	30
Gambar 4. 7 Hasil <i>Filtering</i> Data	30
Gambar 4. 8 <i>Source Code</i> Untuk Menjumlahkan Data Quantity	31
Gambar 4. 9 Sebelum Proses Penjumlahan <i>Field</i> Quantity.....	31
Gambar 4. 10 Setelah Proses Penjumlahan <i>Field</i> Quantity.....	32
Gambar 4. 11 <i>Source Code</i> <i>Field</i> Total_Order	32
Gambar 4. 12 Hasil Proses Pembuatan <i>Field</i> Total_Orders	33
Gambar 4. 13 <i>Source Code</i> Pembersihan Data Akhir.....	33
Gambar 4. 14 <i>Cleaned Dataset</i>	34
Gambar 4. 15 <i>Source Code</i> Import Dataset.....	34
Gambar 4. 16 Output <i>Source Code</i> Import Dataset	35
Gambar 4. 17 Algoritma <i>K-Means++</i>	36
Gambar 4. 18 Algoritma <i>Forgy Initialization</i>	36
Gambar 4. 19 Inisialisasi 3 <i>Centroid</i> Awal Menggunakan <i>K-Means++</i>	37
Gambar 4. 20 Inisialisasi 2 <i>Centroid</i> Awal Menggunakan <i>K-Means++</i>	37
Gambar 4. 21 Inisialisasi 3 <i>Centroid</i> Awal Menggunakan <i>Forgy Initialization</i> ..	38
Gambar 4. 22 Inisialisasi 2 <i>Centroid</i> Awal Menggunakan <i>Forgy Initialization</i> ..	38
Gambar 4. 23 Rumus <i>Euclidean Distance</i>	40

Gambar 4. 24 Rumus <i>K-Means Clustering</i>	41
Gambar 4. 25 Rumus Iterasi	41
Gambar 4. 26 Iterasi Ke-1	42
Gambar 4. 27 Iterasi Ke-50	42
Gambar 4. 28 <i>Source Code</i> Menghitung <i>Silhouette Coefficient</i>	44
Gambar 4. 29 Visualisasi Ploting <i>K-Means++</i> Dengan 3 <i>Cluster</i>	46
Gambar 4. 30 Visualisasi Ploting <i>Forgy Initialization</i> Dengan (3 <i>Cluster</i>)	48
Gambar 4. 31 Visualisasi Ploting <i>K-Means++</i> Dengan 2 <i>Cluster</i>	50
Gambar 4. 32 Visualisasi Ploting <i>Forgy Initialization</i> Dengan 2 <i>Cluster</i>	52

ANALISIS DATA PENJUALAN PADA TOKO BUKU SOCIAL AGENCY BARU AMBARUKMO MENGGUNAKAN METODE *K-MEANS CLUSTERING*

Mohammad Ikhwanud Dawam

19106050050

INTISARI

Toko SAB Ambarukmo merupakan salah satu toko buku yang cukup besar di Yogyakarta dan mempunyai cukup banyak pelanggan. Penelitian ini dilakukan dengan tujuan menganalisis tingkat penjualan setiap produk sehingga diharapkan dapat meningkatkan tingkat penjualan di Toko SAB Ambarukmo. Data penjualan digunakan sebagai sampel dalam penelitian ini dan data tersebut diolah dengan metode *K-Means Clustering*.

Dalam penelitian ini didapatkan beberapa riwayat data penjualan dari bulan Desember 2022 sampai bulan Mei 2023 jika dijumlahkan terdapat 51948 total transaksi yang berlangsung. Dari data yang didapat diolah melalui tahap *preprocessing data* dan dihasilkan sebanyak 7947 data. Setelah itu, data dianalisis dengan metode *K-Means Clustering*.

Proses *Clustering* menghasilkan 4 output dengan metode dan jumlah cluster yang berbeda, yang pertama dengan metode inisialisasi kluster awal *K-Means++* jumlah *cluster* 3, kemudian metode inisialisasi kluster awal *Forgy Initialization* dengan jumlah *cluster* 3, metode inisialisasi kluster awal *K-Means++* jumlah *cluster* 2, kemudian metode inisialisasi kluster awal *Forgy Initialization* dengan jumlah *cluster* 2. Didapatkan hasil terbaik dari 4 output tersebut yaitu untuk metode inisialisasi kluster awal *K-Means++* jumlah *cluster* 2 dengan perolehan *silhouette score* paling mendekati angka 1 yaitu 0.9905861869634374.

Kata Kunci: *Clustering, K-Means, Forgy Initialization, K-Means++*

ANALYSIS OF SALES DATA AT THE NEW SOCIAL AGENCY BOOKSTORE AMBARUKMO USING THE K-MEANS CLUSTERING METHOD

Mohammad Ikhwanud Dawam

19106050050

ABSTRACT

SAB Ambarukmo Store is a prominent bookstore in Yogyakarta with a substantial customer base. This research was conducted with the aim of analyzing the survival rate of each product so that it is expected to increase the sales level in SAB Ambarukmo Store. Sales data was sampled for this study, and K-Means Clustering was used to process the data.

Several historical sales data from December 2022 to May 2023 were collected for this research, if added up, there were 51948 total transactions that took place. After obtaining the data, it went through a preprocessing stage and resulted in 7947 data points. The data is then analyzed using the K-Means Clustering method.

The clustering process generates four outputs using different methods and numbers of clusters. The first output uses the K-Means++ initialization method with three clusters, the second one uses the Forgy Initialization method with three clusters, and the third and fourth ones use the K-Means++ and Forgy Initialization methods, respectively, with two clusters each. The best outcome among the four generated outputs is achieved by using the K-Means++ initialization method with two clusters, obtaining a silhouette score closest to 1 (0.9905861869634374).

Keywords: *Clustering, K-Means, Forgy Initialization, K-Means++*

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah berdampak luas di berbagai sektor, tak terkecuali toko retail seperti toko buku, yang mengakibatkan bertambahnya jumlah data yang dihasilkan dari waktu ke waktu, karena hal tersebut kebutuhan untuk mengelola dan memanfaatkan data secara efisien pun meningkat.

Toko Buku Social Agency Ambarukmo adalah sebuah toko buku yang berlokasi di Yogyakarta dan cukup banyak diminati oleh para pembeli dari kalangan anak-anak sampai dewasa, toko ini banyak menjual berbagai macam jenis buku, mulai dari fiksi, nonfiksi, sastra klasik, hingga buku anak-anak. Seiring dengan pertumbuhan bisnis, toko ini menghasilkan sejumlah besar data penjualan, namun belum dimanfaatkan secara maksimal.

Menganalisis tingkat penjualan suatu produk merupakan suatu langkah penting dalam strategi penjualan. Hal ini membantu dalam mengidentifikasi tren dan pola penjualan. Dengan memahami tingkat penjualan suatu produk, toko dapat menyusun strategi untuk meningkatkan penjualan produk lainnya, mengatur stok barang, dan merencanakan penawaran atau diskon. Hal ini akan membantu toko untuk memaksimalkan keuntungan dan memenuhi kebutuhan konsumen.

Salah satu metode yang dapat digunakan untuk menganalisis tingkat penjualan suatu produk adalah klusterisasi *K-Means* yang memungkinkan terciptanya pengelompokan data barang ke dalam kategori-kategori yang diinginkan. Hal ini berguna untuk memahami preferensi pelanggan dengan lebih baik. Metode *K-Means Clustering* dipilih karena proses dari *clustering* dapat diintegrasikan langsung dengan sistem yang sudah ada pada toko tersebut dalam hal ini toko Social Agency Baru Ambarukmo dan juga dapat menghasilkan penemuan hubungan atau pola yang tidak terduga diantara variabel-variabel yang diuji.

Dalam skripsi ini, penulis akan fokus pada analisis tingkat penjualan produk di Toko Buku Social Agency Ambarukmo dengan menggunakan metode klusterisasi *K-Means*. Tujuan penulis adalah untuk memberikan visualisasi penyebaran data yang dapat membantu toko buku dalam meningkatkan penjualan mereka dengan memahami preferensi pelanggan, mengoptimalkan strategi pemasaran, serta pengelolaan manajemen stok yang lebih optimal. Diharapkan melalui penelitian ini, toko dapat memanfaatkan data penjualan mereka secara lebih efektif dan efisien.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka permasalahan yang ingin diteliti adalah:

1. Bagaimana penerapan algoritma klusterisasi *K-Means* untuk menganalisis data penjualan di Toko Buku Social Agency Baru Ambarukmo?
2. Bagaimana hasil perhitungan dari algoritma klusterisasi *K-Means* dalam menganalisis data penjualan di Toko Buku Social Agency Baru Ambarukmo?
3. Bagaimana perbandingan hasil perhitungan data penjualan antara cluster satu dengan yang lainnya?

1.3 Batasan Penelitian

Dalam rangka melakukan penelitian ini, adapun batasan penelitiannya mencakup:

1. Sumber data berupa excell.
2. Data penjualan yang akan diteliti yaitu data selama 6 bulan (Desember 2022 – Mei 2023).
3. Menggunakan bahasa pemrograman python.
4. Visualisasi hasil analisis data penjualan ditampilkan dalam bentuk visualisasi web yang dibangun menggunakan *framework* streamlit.

5. Parameter yang digunakan untuk analisis data penjualan adalah jumlah keseluruhan order dan jumlah frekuensi barang yang dibeli.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk mencapai beberapa hal, diantara lain:

1. Mengetahui penerapan algoritma klusterisasi *K-Means* untuk menganalisis data penjualan di Toko Buku Social Agency Baru Ambarukmo.
2. Mengetahui hasil perhitungan dari algoritma klusterisasi *K-Means* dalam menganalisis data penjualan di Toko Buku Social Agency Baru Ambarukmo?
3. Mengetahui hasil perbandingan perhitungan data penjualan antara cluster satu dengan yang lainnya.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberi manfaat bagi beberapa pihak, diantaranya yaitu:

1. Penulis/Peneliti

Penelitian ini menjadi salah satu sarana penulis dalam menerapkan pengetahuan di bidang Teknik Informatika yang telah didapatkan selama masa perkuliahan dan menambah wawasan terkait data mining khususnya klusterisasi dengan algoritma *K-Means*, dalam hal ini untuk menganalisis data penjualan.

2. Pihak Toko Buku Social Agency Baru Ambarukmo

Hasil dari penelitian ini dapat dimanfaatkan sebagai pendukung keputusan dalam meningkatkan strategi penjualan. Diharapkan juga mempermudah dalam mengidentifikasi peluang-peluang baru, sehingga dapat meningkatkan kualitas layanan kepada pelanggan.

3. Pihak Universitas

Penelitian ini dapat dijadikan koleksi pustaka di perpustakaan UIN Sunan Kalijaga Yogyakarta khususnya dalam analisis data penjualan menggunakan metode *K-Means clustering*.

1.6 Keaslian Penelitian

Penelitian mengenai *data mining* khususnya klusterisasi telah banyak dilakukan. Namun, penelitian yang secara khusus membahas implementasi klusterisasi dengan melakukan analisis data penjualan di Toko Buku Social Agency Baru Ambarukmo untuk menentukan tingkat penjualan suatu barang dengan algoritma *K-Means* dan bahasa pemrograman Python belum pernah ditemukan oleh peneliti.

1.7 Sistematika Penelitian

Untuk mempermudah pembaca dalam memahami dan mencari informasi yang dibutuhkan maka dibutuhkan sistematika penelitian. Berikut sistematika dari masing-masing bab yang ada pada penelitian ini:

BAB I

PENDAHULUAN

Bab pendahuluan dalam sebuah skripsi merupakan bagian awal yang berfungsi untuk memperkenalkan topik penelitian kepada pembaca. Bab ini memiliki beberapa tujuan penting, dengan strukturnya diawali dengan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian, dan sistematika penulisan.

BAB II

TINJAUAN PUSTAKA DAN LANDASAN TEORI

Tinjauan pustaka dan landasan teori merupakan bagian penting dalam sebuah skripsi. Bagian ini berfungsi untuk menyajikan informasi terkait dengan teori-teori yang relevan, penelitian terkait, dan studi pustaka yang menjadi dasar penelitian yang dilakukan. Dalam hal ini teori berkaitan dengan *clustering*, algoritma *K-Means*, *data mining*, dan hal-hal terkait.

BAB III

METODE PENELITIAN

Bab ketiga menjelaskan tentang metode yang digunakan oleh penulis dalam penelitian ini, yaitu metode klusterisasi menggunakan algoritma *K-Means*.

BAB IV

HASIL DAN PEMBAHASAN

Pada bab yang keempat penulis menjabarkan secara detail proses penelitian dari awal hingga akhir, dan juga menjabarkan hasil analisis dari penelitiannya yaitu tentang analisis data penjualan pada Toko Buku Social Agency Baru Ambarukmo menggunakan metode *K-Means clustering*.

BAB V

PENUTUP

Pada bab terakhir yaitu penutup, terdapat dua bagian, yaitu bagian kesimpulan dan saran. Bagian kesimpulan berisi tentang rangkuman dari hasil penelitian dan bagian saran berisi rekomendasi dari penelitian yang telah dilakukan.

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini bertujuan untuk menerapkan metode klusterisasi pada data penjualan, dengan metode yang digunakan yaitu *K-Means*. Data penjualan yang digunakan meliputi jumlah *quantity* barang dan *total order*. Didapatkan sebanyak 7947 data setelah melewati proses *preprocessing data*.

Berdasarkan analisis data penjualan yang telah dilakukan, *K-Means Clustering* dapat mengkategorikan beberapa produk/barang yang ada di Toko Buku SAB Ambarukmo menjadi 2 dan 3 kategori. Untuk 3 kategori yaitu kurang laris, cukup laris, laris. Untuk 2 kategori yaitu kurang laris dan laris. dan didapatkan untuk kluster terbaik yaitu untuk 2 kategori yang mendapatkan skor *silhouette coefficient* paling mendekati angka 1.

Dari temuan ini dapat diambil manfaat untuk mengembangkan strategi pemasaran yang lebih efektif dan manajemen pengelolaan stok barang yang lebih optimal sehingga dapat meningkatkan kepuasan pelanggan.

5.2 Saran

Pada penelitian ini masih mempunyai beberapa kekurangan. oleh karena itu penulis mempunyai beberapa saran untuk peneliti selanjutnya, yaitu:

1. Ambil data dari beberapa cabang toko yang berbeda untuk analisis. Sehingga dapat meningkatkan hasil keragaman data dan keakuratan informasi. Karena data yang digunakan dalam penelitian ini hanya dari 1 cabang toko SAB saja yaitu cabang Ambarukmo. Sehingga hasil temuan hanya sebatas di toko tersebut saja.
2. Untuk algoritma klusterisasi dapat menggunakan metode klusterisasi yang lainnya seperti DBSCAN dan *Hierarchical Clustering*. Sehingga dapat membandingkan hasil dan memperluas pemahaman tentang *clustering*.

3. Dapat menggunakan bahasa pemrograman lain misalnya bahasa pemrograman R yang didesain untuk analisis statistik dan komputasi. Sehingga dapat membandingkan kecepatan dan keakuratan.
4. Dalam proses analisis data penjualan dapat menggunakan variabel lain diluar variabel *total order* dan *total quantity* yang memungkinkan mendapatkan informasi tersembunyi yang bermanfaat bagi pemilik toko.
5. Selain menganalisis tingkat penjualan per barang dapat juga mencari tingkat penjualan per kategori barang sehingga didapatkan hasil yang lebih beragam.

DAFTAR PUSTAKA

- Aggarwal, C. C., & Reddy, C. K. (2013). *Data Clustering: Algorithms and Applications* (1st ed.). Chapman & Hall/CRC.
- Alim, N. H. (2021). *PENERAPAN K-MEANS CLUSTERING UNTUK PENGELOMPOKAN CLUSTER PASIEN BERDASAR DATA PASIEN SIMPUS*. UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA.
- Arthur, D., & Vassilvitskii, S. (2007). k-means++: the advantages of careful seeding. *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1027–1035.
- Chafid, N., & Wibowo, I. A. (2018). IMPLEMENTASI DATA MINING UNTUK CLUSTERING DAERAH PENYEBARAN PENYAKIT DEMAM BERDARAH DI KOTA TANGERANG SELATAN MENGGUNAKAN ALGORITMA K-MEANS. *Jurnal Satya Informatika*, 3(1), 12–24.
- Dawam, M. I. (2023). *ANALISIS DATA PENJUALAN PADA TOKO BUKU SOCIAL AGENCY BARU AMBARUKMO MENGGUNAKAN METODE K-MEANS CLUSTERING*. Universitas Islam Negeri Sunan Kalijaga.
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix: Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102–109.
- Faisal, M., Zamzami, E. M., & Sutarman. (2020). Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance. *Journal of Physics: Conference Series*, 1566(1), 012112. <https://doi.org/10.1088/1742-6596/1566/1/012112>
- Ferrín, P. F., & Vilela, B. B. (2015). Attitudes and reactions of Galician (Spanish) consumers towards the purchase of products from other regions. *Global Business and Economics Review*, 17(2), 131. <https://doi.org/10.1504/GBER.2015.068563>

- Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, 21, 768–769. <https://api.semanticscholar.org/CorpusID:118110564>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (Third Edition). Elsevier. <https://doi.org/10.1016/C2009-0-61819-5>
- Imron, M., Hasanah, U., & Humaidi, B. (2020). Analysis of Data Mining Using K-Means Clustering Algorithm for Product Grouping. *IJIIS: International Journal of Informatics and Information Systems*, 3(1), 12–22. <https://doi.org/10.47738/ijjis.v3i1.3>
- Indriyani, F., & Irfiani, E. (2019). *Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means (Clustering Sales Data at Outdoor Equipment Stores Using K-Means Method)*. 7(2), 109–113. <https://doi.org/10.30595/juita.v7i2.5529>
- Kharisma, R. B., & Yazid, A. S. (2018). The Mapping of Access Point Workloads at UIN Sunan Kalijaga Based on Log Analysis using K-Means Algorithm. *IJID (International Journal on Informatics for Development)*, 6(1), 17. <https://doi.org/10.14421/ijid.2017.06105>
- Krishna, K., & Narasimha Murty, M. (1999). Genetic K-means algorithm. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 29(3), 433–439. <https://doi.org/10.1109/3477.764879>
- Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118874059>
- Mangku Negara, I. S., Purwono, P., & Ashari, I. A. (2021). Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 dengan Algoritma K-means. *JOINTECS (Journal of Information Technology and Computer Science)*, 6(3), 153. <https://doi.org/10.31328/jointecs.v6i3.2693>
- Muningsih, E., & Kiswati, S. (2015). PENERAPAN METODE K-MEANS UNTUK CLUSTERING PRODUK ONLINE SHOP DALAM PENENTUAN STOK BARANG. *Jurnal Bianglala Informatika*, II(1), 1–10. <https://repository.bsi.ac.id/index.php/unduh/item/235723/Jurnal--Penerapan-Metode-K-Means.pdf>

- Peterson, B. (2014). *General Python FAQ*.
<https://docs.python.org/3/faq/general.html#what-is-python>
- Qomariyah, & Siregar, M. U. (2022). Comparative Study of K-Means Clustering Algorithm and K-Medoids Clustering in Student Data Clustering. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 7(2), 91–99.
<https://doi.org/10.14421/jiska.2022.7.2.91-99>
- Richards, T. (2021). *Getting Started with Streamlit for Data Science: Create and deploy Streamlit web applications from scratch in Python*. Packt Publishing.
- Sani, S. S. (2021). *IMPLEMENTASI ALGORITMA K-MEANS UNTUK CLUSTERING DATA SKRIPSI*.
<https://repositori.usu.ac.id/handle/123456789/47218>
- Sugriyono, S., & Siregar, M. U. (2020). Preprocessing kNN algorithm classification using K-means and distance matrix with students' academic performance dataset. *Jurnal Teknologi Dan Sistem Komputer*, 8(4).
<https://doi.org/10.14710/jtsiskom.2020.13874>
- Wardhani, A. K. (2016). IMPLEMENTASI ALGORITMA K-MEANS UNTUK PENGELOMPOKAN PENYAKIT PASIEN PADA PUSKESMAS KAJEN PEKALONGAN. *Jurnal Transformatika*, 14.
<https://journals.usm.ac.id/index.php/transformatika/article/view/387/248>
- Widiartha, I. M. (2013). PENERAPAN METODE FORGY PADA PERILAKU LEBAH PENJELAJAH DALAM ARTIFICIAL BEE COLONY. *Jurnal Ilmu Komputer*.
- Yonata, Y., Maharani, H., & Viona, C. (2021). Analisis Clustering Pelanggan Berdasarkan Data Transaksi Penjualan Menggunakan Algoritme K-Means dan Metode Recency, Frequency, Monetary (RFM) (Studi Kasus: CV XYZ). *Jurnal Telematika*, 16(2). <https://journal.ithb.ac.id/telematika/article/view/379>