



**SEGMENTASI PELANGGAN AGEN AMI MS GLOW
PETERONGAN MENGGUNAKAN ALGORITMA
K-MODES BERBASIS PYTHON**

SKRIPSI



**MOHAMMAD AINUL YAQIN
4218005**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PESANTREN TINGGI DARUL 'ULUM
JOMBANG
2022**





**SEGMENTASI PELANGGAN AGEN AMI MS GLOW
PETERONGAN MENGGUNAKAN ALGORITMA
K-MODES BERBASIS PYTHON**

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat Memperoleh
Gelar Sarjana pada Program Studi Matematika



MOHAMMAD AINUL YAQIN
4218005

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PESANTREN TINGGI DARUL 'ULUM
JOMBANG
2022**



Hak Cipta Milik Unipdu Jombang

[@www.unipdu.ac.id](http://www.unipdu.ac.id)



HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

SKRIPSI

SEGMENTASI PELANGGAN AGEN AMI MS GLOW PETERONGAN MENGGUNAKAN ALGORITMA *K-MODES* BERBASIS PYTHON

Oleh :

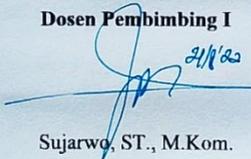
Mochammad Ainul Yaqin
4218005

Telah diperiksa dan disetujui untuk diuji

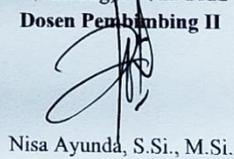
Jombang, 17 Juli 2022

Dosen Pembimbing I

Dosen Pembimbing II

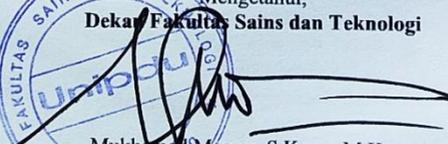


Sujarwo, ST., M.Kom.



Nisa Ayunda, S.Si., M.Si.

Mengetahui,
Dekan Fakultas Sains dan Teknologi



Mukhammad Masrur, S.Kom., M.Kom.





HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

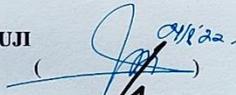
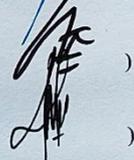
SEGMENTASI PELANGGAN AGEN AMI MS GLOW PETERONGAN MENGGUNAKAN ALGORITMA K-MODES BERBASIS PYTHON

Oleh :

Mochammad Ainul Yaqin
4218005

Telah dipertahankan di depan Tim Penguji
Pada tanggal: 20 Juli 2022

TIM PENGUJI

- 1. Sujarwo, S.T., M.Kom (Ketua) ()
- 2. Wiwit Denny Fitriana, M.Si (Anggota) ()
- 3. Afsah Novita Sari, M.Si (Anggota) ()

Mengetahui,
Dekan Fakultas Sains dan Teknologi



Mukhammad Masrur, S.Kom., M.Kom





PERNYATAAN KEASLIAN TULISAN

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama : Mochammad Ainul Yaqin

NIM : 4218065

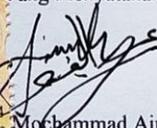
Program Studi : S1 Matematika

Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya dan sungguh-sungguh bahwa skripsi dengan judul **“Segmentasi Pelanggan Agen Ami MS Glow Peterongan Menggunakan Algoritma K-Modes Berbasis Python”** benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil-alihan tulisan atau pikiran orang lain, dan bukan hasil jiplakan.

Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan atau ada yang mengajukan gugatan, maka saya bersedia menerima seluruh sanksi atas perbuatan tersebut, termasuk pembatalan gelar yang saya peroleh dari Universitas Pesantren Tinggi Darul ‘Ulum.

Jombang, 17 Juli 2022

Yang Menyatakan

Mochammad Ainul Yaqin







ABSTRAK

Ms Glow adalah *brand skincare* lokal yang pada tahun 2021 meraih rekor MURI sebagai perusahaan kosmetik dengan jaringan penjualan terbanyak di Indonesia. Salah satu agen resmi Ms Glow di Kabupaten Jombang yang berkembang cukup pesat adalah Agen Ami Ms Glow Peterongan. Meskipun begitu, terdapat dua permasalahan yang dihadapi oleh manajemen yaitu ketidaktepatan stok produk dan penurunan interaksi konten di media sosial. Permasalahan tersebut dapat diatasi apabila manajemen mampu mengidentifikasi karakteristik pelanggannya. Oleh karena itu, pada penelitian ini dilakukan segmentasi pelanggan Agen Ami Ms Glow Peterongan. Metode yang digunakan yaitu Algoritma *K-Modes*. Algoritma ini dapat digunakan untuk menangani data beratribut kategorik. Hal ini sesuai dengan data penelitian yang diolah yang memiliki tipe atribut kategorik. Hasil dari penelitian ini yaitu pelanggan Agen Ami Ms Glow Peterongan terbagi menjadi 3 segmen pelanggan yang didominasi oleh segmen pertama, dengan karakteristik pelanggan: berusia 26 – 35 tahun; tidak memakai KB dan tidak sedang hamil; berjenis kulit berminyak; tidak berjerawat di luar menstruasi; tidak sedang berjerawat; memiliki bekas jerawat dan komedo; tidak berflekk hitam; serta tidak mengalami kulit yang terkadang merah dan gatal, berjumlah 193 pelanggan, segmen kedua berjumlah 76 pelanggan, dan segmen ketiga berjumlah 98 pelanggan.

Kata kunci: *Segmentasi Pelanggan, K-Modes, Python*



ABSTRACT

Ms Glow is a local skincare brand that won the MURI record as a cosmetic company with the most sales network in Indonesia in 2021. One of Ms Glow's official agents in Jombang, that grows rapidly, is Agent Ami Ms Glow Peterongan. However, there are two problems faced by management, namely the inaccuracy of product stock and the drop of content interaction on social media. These problems can be solved if management can identify the characteristics of its customers. Therefore, in this study, customer segmentation of Agent Ami Ms Glow Peterongan was carried out by using the K-Modes Algorithm method. This algorithm can be used to handle categorical attribute data. This is in accordance with the processed research data which has a categorical attribute type. The results of this study are that the customers of Agent Ami Ms Glow Peterongan are divided into 3 segments which are dominated by the first segment, with customer characteristics: aged 26-35 years; not using KB and not pregnant; oily skin type; no acne outside the menstrual cycle; not having acne; have acne scars and comedo; no black spots; and do not experience itchy and reddish skin occasionally, has 193 customers, the second segment has 76 customers, and the third segment has 98 customers.

Keywords: *Customer Segmentation, K-Modes, Python*



KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah Tuhan Yang Maha Esa yang telah memberikan rahmat serta maunah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul, “**Segmentasi Pelanggan Agen Ami Ms Glow Peterongan Menggunakan Algoritma *K-Modes* Berbasis Python**”, ini tepat pada waktunya. Skripsi ini disusun untuk memenuhi salah satu syarat untuk mendapatkan gelar Sarjana Strata 1 (S1) di Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Pesantren Tinggi Darul Ulum Jombang. .

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini, di antaranya yaitu kepada :

1. Kedua orang tua penulis yang senantiasa mendukung dan mendoakan penulis dalam menuntut ilmu di Universitas Pesantren Tinggi Darul ‘Ulum Jombang.



2. Bapak Prof. DR. H. Ahmad Zahro, Lc., MA. selaku rektor Universitas Pesantren Tinggi Darul Ulum Jombang beserta jajarannya yang telah memberikan penulis wadah untuk menimba ilmu.
3. Bapak Muhamad Masrur, S.Kom., M.Kom. selaku Dekan Fakultas Sains dan Teknologi Unipdu,
4. Ibu Afsah Novita Sari, S.Si., M.Si. selaku Kepala Program Studi Matematika,
5. Bapak Sujarwo, S.T., M.Kom. selaku dosen pembimbing satu yang telah ikhlas meluangkan waktu serta dengan penuh kesabaran telah membimbing penulis dalam proses penyusunan skripsi ini.
6. Ibu Nisa Ayunda, S.Si., M.Si. selaku dosen pembimbing dua yang telah ikhlas meluangkan waktu serta dengan penuh kesabaran telah membimbing penulis dalam proses penyusunan skripsi ini.



7. Seluruh dosen Program Studi Matematika dan seluruh staf Fakultas Sains dan Teknologi.

Pemulis menyadari bahwa skripsi ini masih jauh dari sempurna. Masih terdapat banyak kesalahan dan kekurangan baik dari segi materi ataupun cara penyajiannya. Oleh karena itu, penulis mengharapkan saran dan kritik dari semua pihak demi kesempurnaan skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi penulis, pembaca, dan kepada siapa saja yang ingin mengembangkan penelitian ini lebih lanjut.

Jombang, 17 Juli 2022

Penulis



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	iii
HALAMAN PERSETUJUAN	v
HALAMAN PENGESAHAN	vii
PERNYATAAN KEASLIAN TULISAN	ix
ABSTRAK	xi
ABSTRACT	xii
KATA PENGANTAR.....	xiii
DAFTAR ISI	xvi
DAFTAR TABEL	xviii
DAFTAR GAMBAR	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	6
1.3 Tujuan	6
1.4 Manfaat	7
1.5 Batasan Masalah	8
BAB II KAJIAN PUSTAKA	
2.1 Studi Terdahulu.....	11
2.2 Segmentasi Pelanggan	13
2.3 <i>Clustering</i>	15
2.4 Analisis <i>Cluster</i>	17
2.5 Algoritma <i>K-Modes</i>	17
2.6 Python	23
BAB III METODE PENELITIAN	
3.1 Tahapan Penelitian.....	25
3.1.1 Identifikasi Masalah	25
3.1.2 Studi Literatur.....	26
3.1.3 Pengumpulan Data.....	27



3.1.4	<i>Data Cleaning</i>	28
3.1.5	Transformasi Data	28
3.1.6	<i>Clustering</i>	29
3.1.7	Interpretasi	29
3.1.8	Hasil & Kesimpulan	29
3.2	Lokasi dan Waktu Penelitian	30
BAB IV HASIL DAN PEMBAHASAN		
4.1	Hasil	31
4.1.1	Persiapan Data	31
4.1.2	<i>Data Cleaning</i>	32
4.1.3	Transformasi Data	34
4.1.4	<i>Clustering</i>	39
4.1.5	Visualisasi Hasil Cluster	49
4.2	Pembahasan	65
4.2.1	Analisis Hasil Cluster/Segmen	65
BAB V PENUTUP		
5.1	Kesimpulan	71
5.2	Saran	73
DAFTAR PUSTAKA		75
LAMPIRAN		79



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	11
Tabel 3.1 Jadwal Penelitian.....	30



DAFTAR GAMBAR

Gambar 2.1 Diagram Alir Algoritma <i>K-Modes</i>	21
Gambar 3.1 Diagram Alir Tahapan Penelitian	25
Gambar 4.1 Dataset Keluhan Pelanggan	34
Gambar 4.2 Dataset dengan Pengelompokan Usia.....	37
Gambar 4.3 Dataset Tidak Memiliki Data Null.....	38
Gambar 4.4 Proses Perhitungan Nilai Cost K	40
Gambar 4.5 Perhitungan Nilai Cost.....	41
Gambar 4.6 Grafik Perhitungan Nilai Cost	42
Gambar 4.7 Proses <i>K-Modes</i>	43
Gambar 4.8 Proses Penggabungan cluster_predicted dengan Dataset.....	44
Gambar 4.9 Hasil Penggabungan cluster_predicted dengan Dataset.....	46
Gambar 4.10 Informasi Klaster Pertama	48
Gambar 4.11 Informasi Klaster Kedua.....	48
Gambar 4.12 Informasi Klaster Ketiga.....	49
Gambar 4.13 Klaster Berdasarkan Atribut Jerawat di Luar Menstruasi.....	51
Gambar 4.14 Klaster Berdasarkan Atribut Sedang Berjerawat.....	52
Gambar 4.15 Klaster Berdasarkan Atribut Bekas Jerawat.....	54
Gambar 4.16 Klaster Berdasarkan Atribut Flek Hitam	55
Gambar 4.17 Klaster Berdasarkan Atribut Komedo.....	56
Gambar 4.18 Klaster Berdasarkan Atribut Kulit Terkadang Merah dan Gatal	58
Gambar 4.19 Klaster Berdasarkan Atribut Kondisi Kehamilan.....	59



Gambar 4.20 Klaster Berdasarkan Atribut Pemakaian KB	61
Gambar 4.21 Klaster Berdasarkan Atribut Jenis Kulit	62
Gambar 4.22 Klaster Berdasarkan Atribut Usia	64
Gambar 4.23 Proses Menampilkan Pusat Klaster.....	66
Gamabr 4.24 Hasil Pusat Klaster.....	67



DAFTAR LAMPIRAN

Lampiran 1 Formulir Konsultasi Keluhan Pelanggan	79
Lampiran 2 Dataset dengan Penambahan Kolom cluster_predicted.....	80
Lampiran 3 <i>Source Code</i> Program	94





BAB I

PENDAHULUAN

1.1. Latar Belakang

Dewasa ini, industri kecantikan di Indonesia mengalami perkembangan yang cukup pesat. Berdasarkan data yang dirilis oleh Nielsen dan Euromonitor, pertumbuhan penjualan produk kecantikan di Indonesia pada tahun 2017 adalah sebesar 11,99%, yang memberikan kontribusi sebesar USD 1,4 miliar atau setara Rp 19 triliun, nilai ini diperkirakan akan terus meningkat setiap tahunnya. (Supriyatna, I. & Djailani, M. F., 2021). Hal ini tak lepas juga dari banyaknya *brand* kecantikan yang berkembang di Indonesia. Salah satunya yaitu Ms Glow, yang merupakan *brand* kecantikan dan *skincare* lokal yang berdiri sejak tahun 2013. Pada bulan Agustus 2021 yang lalu, Ms Glow meraih rekor MURI sebagai perusahaan kosmetik dengan jaringan penjualan terbanyak di

Indonesia, yakni sebanyak 78.147 jaringan (Warta Ekonomi, 2021). Dalam jaringan penjualannya, Ms Glow dibagi ke dalam 4 tingkatan level, yakni Reseller, Member, Agen, dan Distributor.

Salah satu agen Ms Glow yang berada di Kabupaten Jombang yaitu Ami Ms Glow Peterongan, yang merupakan salah satu agen resmi dari Ms Glow Store yang terletak di Jl. Brawijaya No. 86, Peterongan, Kecamatan Peterongan, Kabupaten Jombang. Agen Ami Ms Glow Peterongan sudah berkembang cukup pesat. Hal ini bisa dilihat dari omset agen Ami Ms Glow Peterongan yang mencapai ratusan juta per bulan dengan jumlah pelanggan yang sudah mencapai ribuan pelanggan. Meskipun begitu, Ami Ms Glow Peterongan memiliki beberapa permasalahan mendasar pada kegiatan usahanya. Berdasarkan wawancara peneliti dengan pemilik agen Ami Ms Glow Peterongan, terdapat dua permasalahan yang sedang dihadapi oleh



manajemen agen. Yang pertama yaitu ketidaktepatan dalam melakukan penyetokan produk. Indikasinya yaitu terdapat beberapa jenis barang yang sepi peminatnya, seperti paket Ms Kids, Ms Slim, *body soap*, dan lain-lain, sehingga membuat barang menjadi lama tidak terjual dan bahkan ada yang tidak terjual. Masalah yang kedua yaitu tingkat interaksi beberapa konten di media sosial menurun. Manajemen Ami Ms Glow Peterongan memanfaatkan media sosial untuk menjalankan strategi pemasaran, yaitu dengan cara membuat konten informatif seputar dunia kecantikan. Tetapi, pada beberapa waktu terakhir, terdapat beberapa konten yang memiliki tingkat interaksi yang kecil. Kedua permasalahan Ami Ms Glow Peterongan tersebut dapat diatasi apabila manajemen agen mampu mengidentifikasi serta memahami tipe dan karakteristik pelanggannya. Apabila karakteristik pelanggan dipahami dengan baik, maka manajemen agen Ami Ms Glow



dapat menyetok barang sesuai dengan kebutuhan pelanggan serta konten informatif di media sosial akan lebih tepat sasaran.

Masalah pemahaman karakteristik pelanggan dapat diselesaikan menggunakan metode segmentasi pelanggan. Menurut Morissan (2010), segmentasi pelanggan adalah upaya pengelompokan pelanggan berdasarkan karakteristik yang sama antar pelanggan. Pelanggan dengan kesamaan tertentu akan tergabung dalam satu segmen, sementara yang berbeda akan tergabung ke dalam segmen yang lain. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan agen Ami Ms Glow Peterongan menggunakan Algoritma *K-Modes*. Hasil dari penelitian ini diharapkan dapat membantu manajemen agen Ami Ms Glow Peterongan dalam mengenali karakteristik pelanggannya, sehingga nantinya akan berguna dalam penerapan strategi



pemasaran dan perencanaan stok barang agen Ami Ms Glow Peterongan.

Metode yang umumnya digunakan untuk menyelesaikan masalah segmentasi yaitu metode *K-Means*, tetapi metode ini hanya bekerja pada *dataset* dengan tipe atribut numerik (Kurniawan, 2020). Pada penelitian ini, data yang akan diolah adalah data isian formulir konsultasi keluhan yang dialami oleh pelanggan, yang mana *dataset* yang diolah tersebut atributnya bertipe kategorik, sehingga dibutuhkan metode yang dapat mengatasi data bertipe kategorik. Algoritma *K-Modes* adalah metode yang dikembangkan dari algoritma *K-Means* yang dapat diterapkan pada data bertipe kategorik. Kelebihan dari algoritma *K-Modes* yaitu dapat menghasilkan kluster yang lebih stabil dengan waktu komputasi yang lebih singkat dibandingkan dengan algoritma *K-Means* (Nduru dkk, 2018).



1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah pada penelitian ini adalah sebagai berikut:

- a. Bagaimana penerapan algoritma *K-Modes* pada segmentasi pelanggan Agen Ami Ms Glow Peterongan?
- b. Bagaimana analisis terhadap hasil segmentasi pelanggan Agen Ami Ms Glow Peterongan?

1.3. Tujuan

Tujuan dari penelitian ini yaitu sebagai berikut:

- a. Untuk mengetahui bagaimana algoritma *K-Modes* diterapkan dalam melakukan segmentasi pelanggan Agen Ami Ms Glow Peterongan.
- b. Menganalisis segmen pelanggan dari hasil segmentasi menggunakan algoritma *K-Modes*.



1.4. Manfaat

Manfaat yang bisa diambil dari penelitian ini adalah:

- a. Sebagai wawasan bagi manajemen Agen Ami Ms Glow Peterongan mengenai karakteristik pelanggannya.
- b. Sebagai dasar untuk membuat konten promosi di media sosial yang sesuai dengan karakteristik masing-masing segmen pelanggan.
- c. Sebagai bahan pertimbangan bagi manajemen agen Ami Ms Glow Peterongan dalam melakukan penyetokan barang jualannya.
- d. Mengaplikasikan ilmu yang didapat di bangku perkuliahan untuk mengatasi permasalahan di dunia nyata.
- e. Menambah wawasan serta pengalaman mahasiswa terkait dengan dunia usaha dan pemasaran.



1.5. Batasan Masalah

Adapun batasan-batasan masalah dalam penelitian ini adalah sebagai berikut:

- a. Data yang diolah adalah data isian formulir konsultasi keluhan yang dialami oleh pelanggan sebelum melakukan pembelian produk di agen Ami Ms Glow Peterongan. Atribut data yang diolah yaitu usia, berjerawat di luar periode menstruasi, ada tidaknya jerawat, bekas jerawat, flek hitam, komedo, jenis kulit, sedang hamil atau tidak, kulit terkadang merah dan gatal atau tidak dan menjalankan program KB atau tidak.
- b. Data yang diolah adalah data formulir pada rentang waktu 1 Januari sampai dengan 28 Februari 2022.
- c. Metode *clustering* yang digunakan yaitu metode *K-Modes*.



- d. Perangkat lunak yang digunakan sebagai alat bantu penyelesaian masalah segmentasi adalah Google Colaboratory Notebook.





Hak Cipta Milik Unipdu Jombang

[@www.unipdu.ac.id](http://www.unipdu.ac.id)



BAB II

KAJIAN PUSTAKA

2.1 Studi Terdahulu

Beberapa penelitian terdahulu yang digunakan sebagai referensi atau dasar dalam melakukan penelitian tugas akhir ini yaitu sebagai berikut:

Tabel 2.1 Penelitian Terdahulu

Judul Penelitian 1	Segmentasi Pelanggan Produk <i>Digital Service</i> Indihome Menggunakan Algoritma K-Means Berbasis Python
Penulis/Tahun	Nisa Hanum Harani, Cahyo Priyanto, & Fikri Adi Nugraha/2020
Deskripsi	Penelitian ini mensegmentasi pelanggan produk <i>digital service</i> Indihome menggunakan metode K-Means. Penentuan jumlah kluster menggunakan metode Elbow dengan $k = 3$. Penelitian ini menghasilkan tiga segmen pelanggan, yaitu segmen pelanggan dengan yang menghasilkan keuntungan terbesar, pelanggan yang berpotensi menghasilkan keuntungan besar, dan pelanggan yang menghasilkan keuntungan lebih sedikit.
Judul Penelitian 2	K-Modes Clustering untuk Mengetahui Jenis Makanan Daerah yang Populer pada Website Resep Online (Studi Kasus: Masakan Banjar di Cookpad.com

Penulis/Tahun	Fatma Indriani & Irwan Budiman/2017
Deskripsi	Penelitian ini membahas tentang pengelompokan pada data resep masakan daerah Banjar dengan algoritma K-Modes untuk mengetahui kelompok masakan yang sejenis berdasarkan bahan masakannya. Hasil penelitian ini yaitu didapat dua alternatif jumlah klaster, $k = 4$ dan $k = 8$. Untuk $k = 4$ menghasilkan kelompok yang lebih umum sedangkan $k = 8$ menghasilkan kelompok yang spesifik. Untuk anggota $k = 4$ terdiri dari sayur asam, soto banjar, masakan gurih lain-lain, kue dan bubur manis. Untuk $K = 8$ anggotanya terdiri dari sayur asam, soto banjar, kue basah, masakan gurih lain-lain, masakan habang, bubur manis, kuah ketupat, dan masakan gurih asam.
Judul Penelitian 3	Pengelompokan Mahasiswa Berbasis Categorical Variables Menggunakan Metode K-Modes Clustering
Penulis/Tahun	Heribertus Yulianton, Felix Andreas Sutanto, & Sri Mulyani/2021
Deskripsi	Penelitian ini melakukan pengelompokan terhadap data mahasiswa Universitas Stikubank menggunakan metode K-Modes. Data mahasiswa yang diolah bertipe kategorik. Hasil dari penelitian ini yaitu mahasiswa terbagi menjadi dua klaster mahasiswa berdasarkan prodi, pekerjaan, dan pendidikan orang tua dengan klaster 1 yang mendominasi secara signifikan.

Berdasarkan beberapa studi terdahulu yang telah dijabarkan pada tabel (2.1) di atas, topik yang dibahas



pada penelitian tugas akhir ini yaitu segmentasi pelanggan agen Ami Ms Glow Peterongan menggunakan algoritma *K-Modes*.

2.2 Segmentasi Pelanggan

Wendell (1956) mengemukakan bahwa segmentasi adalah usaha untuk mengelompokkan pelanggan berdasarkan kebutuhan. Tujuannya yaitu memberikan produk tunggal tertentu kepada sekumpulan pelanggan yang telah diidentifikasi memiliki kebutuhan tertentu, sehingga strategi pemasaran perusahaan dapat diterapkan dengan lebih efektif dan ekonomis. Asumsi yang mendasari segmentasi pelanggan menurut Blythe (2005) yaitu:

- a) Tidak semua karakteristik konsumen mirip satu sama lain.
- b) Subkelompok konsumen dengan latar belakang, perilaku, nilai, dan kebutuhan sama dapat dikenali.



- c) Subkelompok lebih kecil dan lebih homogen dibandingkan pasar secara keseluruhan.
- d) Lebih mudah memuaskan sekumpulan kecil konsumen yang sama dibanding berupaya memuaskan kumpulan besar konsumen yang beragam.

Berikut adalah pentingnya melakukan segmentasi pelanggan menurut Hooley, Piercy dan Nicolaud (2008):

- a) Memungkinkan perusahaan kecil untuk menciptakan ceruk pasar dan mempertahankannya karena mampu menyelaraskan kompetensinya dengan sasaran pasar.
- b) Mampu mengidentifikasi celah atau gap di pasar, yaitu segmen yang selama ini belum atau tidak terlayani secara baik.
- c) Mampu mengidentifikasi segmen yang sudah menurun di pasar.



- d) Memungkinkan perusahaan untuk menyesuaikan produk atau jasanya dengan kebutuhan pasar.
- e) Perusahaan akan kehilangan keunggulan kompetitif apabila tidak melakukan segmentasi.

Menurut Tjiptono (2015), pada intinya, segmentasi pasar adalah proses pengelompokan pasar sebuah produk atau jasa yang heterogen ke dalam beberapa segmen, yang mana tiap segmen cenderung mempunyai kesamaan dalam hal kebutuhan, keinginan, perilaku, serta respon terhadap program pemasaran.

2.3 Clustering

Clustering atau klasterisasi adalah salah satu metode pada data mining yang tujuannya untuk mengelompokkan data atau objek ke dalam kelompok-kelompok. *Cluster* adalah kumpulan objek data yang memiliki kesamaan karakteristik antar data dalam satu *cluster* dan memiliki ketidaksamaan karakteristik terhadap objek data yang berbeda *cluster* (Rohmawati N., S.



Defiyanti, dan M. Jajuli, 2015). *Clustering* merupakan pembelajaran yang tidak terawasi dimana tidak memerlukan target luaran atau yang sering disebut *unsupervised learning* (Prasetyo E., 2014).

Terdapat dua metode dalam *Clustering*, yaitu metode hirarki dan metode nonhirarki. Metode hirarki adalah metode yang dimulai dengan mengelompokkan beberapa objek yang memiliki kesamaan paling dekat. Kemudian diteruskan ke objek lain yang memiliki kedekatan lainnya. Sehingga, *cluster* akan membentuk hirarki dari yang paling mirip sampai yang tidak mirip. Sedangkan metode nonhirarki dimulai dengan menentukan terlebih dahulu jumlah *cluster* baru dilakukan proses *clustering*. Menurut Han dkk. (2006), metode hirarki digunakan apabila jumlah klaster belum diketahui, sedangkan pada metode nonhirarki bertujuan mengelompokkan objek ke ke dalam k buah *cluster*.



2.4 Analisis Cluster

Hair dkk (2014) mengemukakan bahwa analisis kluster adalah kumpulan dari beberapa teknik pengolahan data multivariabel yang bertujuan untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Hasil dari kluster yang terbentuk harus menunjukkan kesamaan internal yang tinggi dalam satu kluster dan keragaman yang tinggi antar kluster. Prinsip dari analisis kluster adalah proses untuk mereduksi sekumpulan objek besar menjadi lebih kecil yang disebut kluster (Fahmi M.F., 2016).

2.5 Algoritma K-Modes

Algoritma *K-Modes* diperkenalkan oleh Huang pada tahun 1997. Algoritma ini merupakan hasil dari modifikasi algoritma *clustering K-Means* yang mana hanya bisa digunakan untuk mengelompokkan data numerik, sedangkan algoritma *K-Modes* dapat digunakan untuk mengelompokkan data kategorik.



Data dengan karakteristik yang sama akan bergabung ke dalam satu kelompok dan yang berbeda akan bergabung ke dalam kelompok yang lain. Tujuan algoritma ini untuk meminimalkan variasi dalam satu kelompok dan memaksimalkan variasi antar kelompok. Kelebihan dari algoritma *K-Modes* yaitu dapat menghasilkan kluster yang lebih stabil dengan waktu komputasi yang lebih singkat dibandingkan dengan algoritma *K-Means* (Nduru dkk, 2018). Karena merupakan pengembangan dari algoritma *K-Means*, algoritma ini termasuk ke dalam metode *clustering* nonhirarki.

Prasetyo E. (2014) menyebutkan beberapa modifikasi yang dilakukan terhadap algoritma *K-Means* adalah sebagai berikut:

- a) Menggunakan ukuran pencocokan ketidaksamaan sederhana dari fitur data berjenis kategorikal.



- b) Menggunakan modus (nilai yang paling sering muncul) sebagai ganti mean klaster.
- c) Untuk mencari modus dari sekumpulan nilai, digunakan metode berbasis frekuensi.

Misal x dan y adalah dua data bertipe kategorik, x sebagai pusat *cluster* dan y sebagai data yang akan dibandingkan dengan pusat cluster. Ukuran ketidaksamaan antara x dan y bisa diukur dengan jumlah ketidakcocokan nilai dari fitur yang berkorespondensi dari dua data. Data tersebut semakin mirip jika semakin kecil nilai ketidakcocokannya. Metrik seperti ini biasa disebut dengan pencocokan sederhana (*simple matching*). Rumus untuk menghitung jarak dalam *K-Modes* adalah sebagai berikut (Yulianita & Istiawan, 2017),

$$d(X, Y) = \sum_{j=1}^r \epsilon(x_j, y_j) \quad \dots\dots\dots (2.1)$$

Di mana :

$$d(x, y) = \text{Jarak data } x \text{ ke } y$$



x_j = Nilai fitur ke- j dari x

y_j = Nilai fitur ke- j dari y

r = Jumlah fitur

Dengan :

$$\epsilon(x_j, y_j) = \begin{pmatrix} 0 & (x_j = y_j) \\ 1 & (x_j \neq y_j) \end{pmatrix}$$

Langkah-langkah dalam melakukan klusterisasi dengan algoritma *K-Modes* sebagai berikut (Prasetyo, 2014):

- a) Memilih k modus awal sebagai titik pusat, satu untuk setiap kluster.
- b) Menghitung jarak masing-masing data terhadap semua titik pusat kluster. Alokasikan setiap objek ke kluster terdekat menggunakan persamaan (2.1).
- c) Memperbarui modus (sebagai pusat kluster baru) dari setiap kluster dengan nilai kategori yang sering muncul pada setiap kluster.
- d) Mengulangi langkah b dan c apabila tidak terdapat data yang berpindah kluster atau posisi pusat kluster tidak berubah



Berikut gambar diagram alir untuk algoritma *K-Modes* (Nduru dkk, 2018).



Gambar 2.1 Diagram alir algoritma *K-Modes*

Keterangan :

- 1) Mulai
- 2) Pre-Processing Data
Melakukan pembersihan dan transformasi data.
- 3) Menentukan Jumlah *Cluster*
Memilih k kluster dan data secara acak untuk menentukan pusat *cluster* (centroid).
- 4) Menentukan centroid secara acak
Menginisialisasi dan iterasi data secara acak/mengambil data yang digunakan sebagai centroid awal.
- 5) Menghitung jarak setiap data pada centroid
Menghitung ketidakmiripan masing-masing data ke centroid menggunakan persamaan (2.1).
- 6) Kelompokkan data berdasarkan jarak terkecil
Mengalokasikan data ke kluster berdasarkan jarak terkecil yang di dapat dari penghitungan jarak pada langkah sebelumnya.



- 7) Hitung kembali pusat centroid yang terbentuk
Mendapatkan pusat *cluster* baru setelah pengalokasian data pada langkah sebelumnya.
- 8) Centroid berubah?
Jika centroid berubah, maka dilakukan iterasi kembali dari langkah menghitung jarak setiap data pada centroid. Jika centroid tidak berubah, maka iterasi selesai/proses *clustering* selesai.
- 9) Selesai.

2.6 Python

Python merupakan salah satu bahasa pemrograman terpopuler di dunia (Kumar, 2018). Python banyak digunakan dalam bidang analisis data, data mining, *Machine Learning* dan kecerdasan buatan. Popularitas Python ini salah satunya adalah tersedianya banyak *library* (modul siap pakai) yang mendukung kegiatan analisis data dan *Machine Learning*, seperti Matplotlib, Numpy, Pandas, Scikit-learn, dan *library* lainnya.



Referensi kode program untuk implementasi algoritma *K-Modes* yang digunakan pada penelitian ini merujuk pada kode program yang ditulis oleh Ashish (2019) dalam situs Kaggle (sebuah situs komunitas *data science*) mengenai implementasi *K-Modes*.





BAB III METODE PENELITIAN

3.1 Tahapan Penelitian

Adapun tahapan dalam penelitian ini dapat dilihat pada gambar diagram alir berikut.



Gambar 3.1 Diagram alir tahapan penelitian

3.1.1. Identifikasi Masalah

Tahap awal dari penelitian ini adalah identifikasi masalah. Yaitu melakukan analisis terhadap permasalahan yang akan dijadikan topik penelitian ini. Penelitian ini mengambil studi kasus pada Agen Ami Ms Glow Peterongan

dengan permasalahan yang diangkat yaitu segmentasi pelanggan agen Ami Ms Glow Peterongan berdasarkan keluhan yang dialami pelanggan.

3.1.2. Studi Literatur

Pada tahap ini dilakukan penggalian informasi yang dapat dijadikan referensi dalam pelaksanaan penelitian. Sumber informasi didapatkan dari narasumber atau pemilik agen Ami Ms Glow Peterongan, buku, buku elektronik, penelitian terdahulu seperti pada *paper* dan jurnal, tugas akhir, artikel atau dokumen lainnya yang terkait dengan segmentasi pelanggan. Referensi yang diperlukan yaitu yang berhubungan dengan segmentasi pelanggan, *clustering*, dan algoritma *K-Modes*.



3.1.3. Pengumpulan Data

Setelah mengidentifikasi permasalahan dan mendapatkan semua sumber informasi yang diperlukan, tahap selanjutnya yaitu mengumpulkan data-data yang dibutuhkan sebagai pendukung utama dalam pelaksanaan penelitian ini. Jenis data yang digunakan dalam penelitian ini yaitu data primer, berupa data isian formulir konsultasi keluhan pelanggan agen Ami MS Glow Peterongan yang diambil pada rentang waktu 1 Januari hingga 28 Februari 2022. Sumber data didapatkan dari manajemen Agen Ami MS Glow Peterongan. Instrumen penelitian yang digunakan adalah formulir konsultasi keluhan yang diisi oleh pelanggan sebelum melakukan pembelian produk untuk yang pertama kalinya di Agen Ami MS Glow



Peterongan. Formulir konsultasi keluhan pelanggan dapat dilihat pada lampiran 1.

3.1.4. Data Cleaning

Pada tahap ini, dilakukan pembersihan terhadap data, yakni menghilangkan beberapa kolom yang tidak diperlukan untuk proses olah data. Sehingga *dataset* hanya akan berisi atribut yang dibutuhkan untuk proses olah data, yaitu usia, berjerawat di luar periode menstruasi, ada tidaknya jerawat, bekas jerawat, flek hitam, komedo, jenis kulit, sedang hamil atau tidak, kulit terkadang merah dan gatal atau tidak, dan menjalankan program KB atau tidak.

3.1.5. Transformasi Data

Sebelum melakukan proses *clustering*, terlebih dahulu harus melakukan transformasi data. Dari data isian formulir keluhan pelanggan, data umur pelanggan adalah data yang perlu dilakukan



transformasi, yakni mengubah umur ke dalam kategori-kategori umur.

3.1.6. Clustering

Setelah dilakukan transformasi pada, data sudah siap digunakan untuk proses pengolahan data, yaitu proses *clustering*. Pada penelitian ini, algoritma *clustering* yang digunakan adalah algoritma *K-Modes*. Proses implementasi algoritma *K-Modes* menggunakan bahasa pemrograman Python yang dijalankan di *platform* Google Colaboratory Notebook.

3.1.7. Interpretasi

Tahap selanjutnya yaitu melakukan analisis terhadap pola yang dihasilkan dari proses *clustering* menggunakan algoritma *K-Modes*.

3.1.8. Hasil & Kesimpulan

Tahap terakhir pada penelitian ini yaitu mendapatkan hasil segmentasi pelanggan dan



penarikan kesimpulan dari kesemua proses yang telah dilaksanakan dalam penelitian dari awal hingga akhir.

3.2 Lokasi dan Waktu Penelitian

Lokasi penelitian ini yaitu di Agen Ami MS Glow Peterongan. Penelitian ini dilaksanakan pada bulan November 2021 sampai dengan bulan Mei 2022. Berikut adalah tabel jadwal penelitian:

Tabel 3.1 Jadwal Penelitian

Tahap Penelitian	Bulan						
	Nov	Des	Jan	Feb	Mar	Apr	Mei
Identifikasi Masalah							
Studi Literatur							
Pengumpulan Data							
Data Cleaning							
Transformasi Data							
Clustering							
Interpretasi							
Hasil & Kesimpulan							
Penyusunan Laporan							

