

**KLASTERISASI PELANGGAN MENGGUNAKAN  
*ALGORITMA K-MEANS CLUSTERING*  
PADA TOP 5 *ONLINE MARKETPLACE* DI INDONESIA**

**TESIS**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar  
Magister Manajemen**



**UNIVERSITAS  
BAKRIE**

**ANDARA RAHMA DANURISA  
2201021004**


**PROGRAM STUDI  
MAGISTER MANAJEMEN  
UNIVERSITAS BAKRIE  
JAKARTA  
TAHUN 2022**

## HALAMAN PERNYATAAN ORISINALITAS

**Tesis ini adalah hasil karya saya sendiri,  
dan semua sumber baik yang dikutip maupun dirujuk  
telah saya nyatakan dengan benar**

**Nama : Andara Rahma Danurisa**

**NIM : 2201021004**

**Tanda Tangan :** 

**Tanggal : 17 Agustus 2022**


## HALAMAN PENGESAHAN

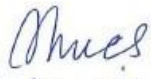
Tesis ini diajukan oleh :

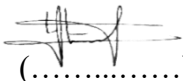
Nama : Andara Rahma Danurisa  
NIM : 2201021004  
Program Studi : Magister Manajemen  
Judul Tesis : Klasterisasi Pelanggan Menggunakan Algoritma  
*K-Means Clustering* Pada *Top 5 Online Marketplace*  
di Indonesia

**Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Manajemen pada Program Studi Magister Manajemen - Universitas Bakrie**

## DEWAN PENGUJI

Pembimbing : Dr. Jerry Heikal, ST, MM   
(.....)

Penguji 1 : Muchsin Saggaff Shihab, SE., MSc., MBA., PhD   
(.....)

Penguji 2 : Arief Bimantoro Suharko, Ph.D   
(.....)

Ditetapkan di : Jakarta

Tanggal : 17 Agustus 2022

## UNGKAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tesis ini. Penulisan Tesis ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Magister Manajemen Program Studi Magister Manajemen pada Fakultas Ekonomi dan Ilmu Sosial Universitas Bakrie. Penulis percaya bahwa keberhasilan penulisan dalam menyelesaikan Tesis ini adalah berkat dari Allah SWT serta tidak terlepas dari bantuan semua pihak mulai dari masa perkuliahan sampai dengan akhir perkuliahan. Penulis berterimakasih kepada semua pihak yang telah memberikan kontribusi dalam penyelesaian Tesis ini dan secara khusus pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada yang terhormat:

- 1) Dr. Jerry Heikal, ST, MM, selaku dosen pembimbing yang selalu memberikan bimbingan, kepercayaan, dan dukungan serta senantiasa ikhlas menyediakan waktu, tenaga, dan pikiran untuk mengarahkan penulis dalam penyusunan Tesis ini;
- 2) Muchsin Saggaff Shihab, SE., MSc., MBA., PhD, dan Arief Bimantoro Suharko, Ph.D, selaku dosen penguji ahli yang telah berkenan untuk menguji, memberikan masukan dan saran kepada penulis sehingga penulis dapat menyelesaikan Tesis ini;
- 3) Seluruh staf dan pengajar Program Studi Magister Manajemen yang telah memberi bantuan, ilmu dan bimbingan;
- 4) Seluruh responden dengan berbagai kriteria penetapan yang telah bersedia menjadi responden dan membantu penulis dalam memperoleh data yang diperlukan;
- 5) Kedua orang tua tercinta dan keluarga besar SAMAWA yang tiada hentinya mendoakan, mendukung, dan memberikan motivasi hingga penulis dapat mencapai titik ini;
- 6) Almarhum Laksma TNI (Purn) dr. H. Subagyo, Sp.B, Sp.OT, FICS dan Ny. Pruning Subagyo selaku Papa dan Mama mertua yang telah memberikan bantuan dukungan material dan moral, sekaligus sebagai sponsor utama bagi penulis dalam melanjutkan jenjang S-2 ini;

- 7) Suamiku tercinta, Mayor Laut (K) dr. Indra Widjayanto, M.Kes, Sp.B yang selalu memberikan dukungan berupa semangat, restu, dan doa, serta bersedia dengan ikhlas mem-*backup* segala hal yang penulis butuhkan;
- 8) Anak-anakku tersayang (KaZaNea) dan anakku kelak kapan kamu akan datang, Insya Allah ilmu ini kelak akan bermanfaat untuk mendidikmu;
- 9) Para sahabat terbaik penulis dalam jenjang S-2 yaitu Nugrahanti Khairani Aisyah, Ajeng Damar Rarasati, M. Cholilurrahman, dan Rudhy Hendarto yang sudah bersama sejak tugas pertama hingga tugas terakhir Tesis. Penulis bersyukur dipertemukan dengan kalian karena ambisi yang besar dan saling suportif bisa mengantarkan kita lulus 1,5 tahun.
- 10) Seluruh teman-teman *Batch* 15, terima kasih atas rasa kekeluargaan yang sangat tulus dalam pertemanan selama masa perkuliahan ini;
- 11) Serta rekan, kerabat, dan sahabat yang tidak bisa disebutkan satu persatu yang telah memberi semangat dan dukungan selama menyelesaikan Tesis ini.
- 12) *Last but not least, I wanna thank me for believing in me, for doing all this hard work, for having no days off, for never quitting, for always being a giver and tryna give more than I receive, for tryna do more right than wrong, for just being me at all times.*

Akhir kata, penulis berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Tesis ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 17 Agustus 2022

Penulis/ Peneliti,



Andara Rahma Danurisa

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Bakrie, saya yang bertanda tangan dibawah ini:

Nama : Andara Rahma Danurisa

NIM : 2201021004

Program Studi : Magister Manajemen

Jenis Tugas Akhir : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty, Free Right*), atas karya ilmiah saya yang berjudul:

### **KLASTERISASI PELANGGAN MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA TOP 5 ONLINE MARKETPLACE DI INDONESIA**

Beserta perangkat yang ada (jika diperlukan) dengan Hak Bebas Royalti Noneksklusif ini Universitas Bakrie berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan Tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 17 Agustus 2022

Yang menyatakan



(Andara Rahma Danurisa)

**KLASTERISASI PELANGGAN MENGGUNAKAN ALGORITMA  
K-MEANS CLUSTERING PADA TOP 5 ONLINE MARKETPLACE  
DI INDONESIA**

Andara Rahma Danurisa

---

**ABSTRAK**

Tokopedia, Shopee, Lazada, Bukalapak dan Orami merupakan 5 marketplace di Indonesia yang memfokuskan pada aktivitas berbelanja menggunakan aplikasi. Jumlah kunjungan pelanggan sangat fluktuatif sehingga mengalami kesulitan untuk menentukan minat pelanggan dalam pembelian suatu produk. Tujuan dari penelitian ini adalah untuk mengidentifikasi karakteristik, mengetahui kategori produk dan kategori *merchant* pada top 5 *online marketplace* di Indonesia menggunakan algoritma *K-Means Clustering*. Terdapat tiga variabel yakni karakteristik pelanggan, kategori produk, dan kategori *merchant*. Pengolahan data dibantu dengan aplikasi SPSS V.25. Pelanggan setia *marketplace* Tokopedia adalah cluster 10. Selain itu, *marketplace* Tokopedia juga memiliki pelanggan yang loyal pada cluster 5. Pelanggan setia *marketplace* Shopee adalah cluster 8. Lalu, *marketplace* Shopee juga memiliki pelanggan yang loyal pada 6 cluster lainnya, diantaranya yaitu pada cluster 2, cluster 3, cluster 4, cluster 6, cluster 7, dan cluster 9. Pelanggan setia *marketplace* Lazada adalah cluster 6. Namun pada cluster 6 tersebut, *marketplace* Shopee masih unggul sebagai *marketplace* yang sering digunakan pada cluster 6. Pelanggan setia *marketplace* Bukalapak adalah cluster 7. Namun pada cluster 7, *marketplace* Shopee masih unggul sebagai *marketplace* yang sering digunakan pada cluster 7. Pelanggan setia marketplace Orami adalah cluster 1.

Kata kunci: *Marketplace*, *Data Mining*, Klasterisasi Pelanggan, *Clustering*, *K-Means*.

***CUSTOMER CLUSTERIZATION USING ALGORITHM  
K-MEANS CLUSTERING ON TOP 5 ONLINE MARKETPLACE  
IN INDONESIA***

Andara Rahma Danurisa

---

***ABSTRACT***

*Tokopedia, Shopee, Lazada, Bukalapak and Orami are 5 marketplaces in Indonesia that focus on shopping activities using applications. The number of customer visits is very volatile so it is difficult to determine customer interest in purchasing a product. The purpose of this study is to identify the characteristics, product categories and merchant in the top 5 online marketplaces in Indonesia using the K-Means Clustering algorithm. There are three variables, namely customer characteristics, product categories, and merchant. Data processing is assisted by the application of SPSS V.25. Loyal customers of Tokopedia are cluster 10. In addition, Tokopedia also has loyal customers in cluster 5. Loyal customers of Shopee are cluster 8. Then, Shopee also has loyal customers in 6 clusters, including cluster 2, cluster 3, cluster 4, cluster 6, cluster 7, and cluster 9. Loyal customers of Lazada are cluster 6. However, in cluster 6, Shopee still excels as marketplace that is often used in cluster 6. The loyal customers of Bukalapak are cluster 7. However, in cluster 7, Shopee still excels as marketplace that is often used in cluster 7. And loyal customers of Orami are cluster 1.*

*Keywords: Marketplace, Data Mining, Customer Clusterization, Clustering, K-Means.*



## DAFTAR ISI

|   |             |
|---|-------------|
| <b>HALAMAN JUDUL .....</b>  | <b>i</b>    |
| <b>HALAMAN PERNYATAAN ORISINALITAS .....</b>                            | <b>ii</b>   |
| <b>HALAMAN PENGESAHAN.....</b>  | <b>iii</b>  |
| <b>UNGKAPAN TERIMA KASIH .....</b>                                      | <b>iv</b>   |
| <b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....</b>                    | <b>vi</b>   |
| <b>ABSTRAK .....</b>  | <b>vii</b>  |
| <b>ABSTRACT .....</b>   | <b>viii</b> |
| <b>DAFTAR ISI.....</b>  | <b>ix</b>   |
| <b>DAFTAR GAMBAR.....</b>   | <b>xii</b>  |
| <b>DAFTAR TABEL .....</b>   | <b>xiii</b> |
| <b>DAFTAR LAMPIRAN .....</b>  | <b>xiv</b>  |
| <br>  |             |
| <b>BAB I PENDAHULUAN .....</b>  | <b>1</b>    |
| 1.1 Latar Belakang .....  | 1           |
| 1.2 Rumusan Masalah .....   | 7           |
| 1.3 Tujuan Penelitian .....   | 7           |
| 1.4 Manfaat Penelitian .....  | 8           |
| <br>  |             |
| <b>BAB II KAJIAN PUSTAKA, KERANGKA PEMIKIRAN DAN<br/>HIPOTESIS.....</b> | <b>9</b>    |
| 2.1 Kajian Pustaka.....   | 9           |
| 2.1.1 <i>Clustering/</i> Klasterisasi .....                             | 9           |
| 2.1.2 Segmentasi Pelanggan.....   | 10          |
| 2.1.3 Tipe Segmentasi Pelanggan .....                                   | 10          |
| 2.1.4 <i>Marketplace</i> .....  | 12          |
| 2.1.5 Algoritma <i>K-Means</i> .....                                    | 12          |
| 2.1.6 Definisi Data Mining.....   | 13          |
| 2.1.7 Top 5 <i>Online Marketplace</i> di Indonesia.....                 | 16          |
| 2.1.8 Produk .....  | 18          |
| 2.2 Penelitian Terdahulu .....  | 22          |
| 2.3 Kerangka Pemikiran.....   | 35          |

|   |           |
|---|-----------|
| <b>BAB III METODE PENELITIAN.....</b>                     | <b>36</b> |
| 3.1 Tahapan Penelitian .....                              | 36        |
| 3.2 Variabel Penelitian .....                             | 37        |
| 3.3 Definisi Operasional Variabel.....                    | 37        |
| 3.4 Populasi dan Sampel .....                             | 41        |
| 3.4.1 Populasi .....                                      | 41        |
| 3.4.2 Teknik Pengambilan Sampel.....                      | 42        |
| 3.5 Penentuan Informan .....                              | 43        |
| 3.5.1 Informan .....                                      | 43        |
| 3.5.2 Teknik Penentuan Informan .....                     | 43        |
| 3.6 Analisis <i>K-Means Clustering</i> .....              | 44        |
| 3.7 Alur Penelitian .....                                 | 45        |
| 3.8 Arti Simbol <i>Flowchart</i> .....                    | 46        |
| 3.9 Jenis dan Sumber Data .....                           | 47        |
| 3.9.1 Data Primer .....                                   | 47        |
| 3.9.2 Data Sekunder .....                                 | 47        |
| 3.10 Teknik Pengumpulan Data.....                         | 47        |
| <b>BAB IV HASIL DAN PEMBAHASAN.....</b>                   | <b>49</b> |
| 4.1 Hasil Penelitian .....                                | 49        |
| 4.1.1 Karakteristik Responden .....                       | 49        |
| 4.1.2 <i>Marketplace</i> yang dipilih Responden.....      | 54        |
| 4.1.3 Pengumpulan Data .....                              | 54        |
| 4.1.4 Penyeleksian Data .....                             | 55        |
| 4.1.5 <i>Pre-Processing Data</i> .....                    | 56        |
| 4.1.6 Transformasi Data .....                             | 56        |
| 4.1.7 Pengolahan Data.....                                | 57        |
| 4.1.8 Analisis <i>Cluster</i> dengan <i>K-Means</i> ..... | 57        |
| 4.2 Pembahasan.....                                       | 121       |
| 4.2.1 <i>Cluster 1</i> .....                              | 121       |
| 4.2.2 <i>Cluster 2</i> .....                              | 122       |
| 4.2.3 <i>Cluster 3</i> .....                              | 123       |
| 4.2.4 <i>Cluster 4</i> .....                              | 124       |

|                       |                         |            |
|-----------------------|-------------------------|------------|
| 4.2.5                 | <i>Cluster 5</i> .....  | 125        |
| 4.2.6                 | <i>Cluster 6</i> .....  | 127        |
| 4.2.7                 | <i>Cluster 7</i> .....  | 128        |
| 4.2.8                 | <i>Cluster 8</i> .....  | 129        |
| 4.2.9                 | <i>Cluster 9</i> .....  | 130        |
| 4.2.10                | <i>Cluster 10</i> ..... | 131        |
| <b>BAB V</b>          | <b>PENUTUP</b> .....    | <b>132</b> |
| 5.1                   | Kesimpulan .....        | 132        |
| 5.2                   | Saran.....              | 136        |
| <b>DAFTAR PUSTAKA</b> | .....                   | <b>138</b> |

## DAFTAR GAMBAR

|             |  |     |
|-------------|--|-----|
| Gambar 1.1  | Top 5 Online Marketplace di Indonesia Kuartal 4 Tahun 2021 .....   | 2   |
| Gambar 1.2  | Top 5 Online Marketplace di Indonesia Kuartal 3 Tahun 2021 .....   | 3   |
| Gambar 2.1  | Kerangka Pemikiran .....   | 35  |
| Gambar 3.1  | Tahapan Penelitian .....   | 37  |
| Gambar 3.2  | Flowchart Algoritma K-Means Clustering .....                       | 45  |
| Gambar 4.1  | Data Pelanggan Marketplace.....                                    | 55  |
| Gambar 4.2  | Pengumpulan Data.....  | 55  |
| Gambar 4.3  | Filter Data .....  | 56  |
| Gambar 4.4  | Transformasi Data .....  | 57  |
| Gambar 4.5  | Cluster 1 Food Lovers .....  | 122 |
| Gambar 4.6  | Cluster 2 An Extravagant and Colorful in Citayam Fashion Week..... | 123 |
| Gambar 4.7  | Cluster 3 Beauty Living .....                                      | 124 |
| Gambar 4.8  | Cluster 4 The Passionate Young Mom .....                           | 125 |
| Gambar 4.9  | Cluster 5 Manly Digital Savvy .....                                | 126 |
| Gambar 4.10 | Cluster 6 Budget Saving Young Mom.....                             | 127 |
| Gambar 4.11 | Cluster 7 Chick Urban Teenagers .....                              | 128 |
| Gambar 4.12 | Cluster 8 Millenial TikTok Boy .....                               | 129 |
| Gambar 4.13 | Cluster 9 Adventure Man and The Counting Mountain .....            | 130 |
| Gambar 4.14 | Cluster 10 Sugar Daddy .....                                       | 131 |

## DAFTAR TABEL

|           |  |     |
|-----------|--|-----|
| Tabel 1.1 | Gap Penelitian Tentang Klasterisasi <i>K-Means</i> .....   | 7   |
| Tabel 2.1 | Hasil Penelitian Terdahulu .....   | 22  |
| Tabel 3.1 | Definisi Operasional Variabel.....   | 38  |
| Tabel 3.2 | <i>User Login</i> Pada <i>Top 5 Online Marketplace</i> di Indonesia<br>Berdasarkan Kuartal Q4 2021 ..... | 42  |
| Tabel 3.3 | Arti Simbol <i>Flowchart</i> .....   | 46  |
| Tabel 4.1 | Karakteristik Responden .....  | 49  |
| Tabel 4.2 | <i>Marketplace</i> Pilihan Responden.....  | 54  |
| Tabel 4.3 | <i>Initial Cluster Centers</i> .....   | 58  |
| Tabel 4.4 | <i>Iteration History</i> .....   | 61  |
| Tabel 4.5 | <i>Final Cluster Centers</i> .....   | 62  |
| Tabel 4.6 | <i>Number of Cases in each Cluster</i> .....   | 73  |
| Tabel 4.7 | Provinsi di Setiap <i>Cluster</i> .....  | 74  |
| Tabel 4.8 | <i>Cluster Membership</i> .....  | 76  |
| Tabel 5.1 | Kesimpulan Hubungan Variabel dengan <i>Top 5 Marketplace</i> .....                                       | 132 |

## DAFTAR LAMPIRAN

|            |  |     |
|------------|--|-----|
| Lampiran 1 | Surat Permohonan Kuesioner Penelitian (Angket) |     |
|            | Untuk Responden .....                          | 143 |
| Lampiran 2 | Karakteristik Responden .....                  | 144 |
| Lampiran 3 | Petunjuk Pengisian Kuesioner .....             | 145 |
| Lampiran 4 | Identitas Responden/ Demografi .....           | 150 |
| Lampiran 5 | Kuesioner.....                                 | 146 |
| Lampiran 6 | <i>K-Means Clustering</i> .....                | 148 |
| Lampiran 7 | Hasil Uji Turnitin .....                       | 160 |