

**ANALISIS DATA WISATAWAN MENGGUNAKAN METODE
K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI
PEMASARAN DI MALANG GATEWAY**

SKRIPSI

**Untuk Memenuhi Persyaratan
Memperoleh Gelar Sarjana**

Oleh:

**Marcellina Agatha
NIM 21.51.0001**



PROGRAM STUDI S-1 SISTEM INFORMASI

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
PPKIA PRADNYA PARAMITA
MALANG
2025**

LEMBAR PERSETUJUAN

ANALISIS DATA WISATAWAN MENGGUNAKAN METODE *K-MEANS CLUSTERING* UNTUK MENENTUKAN STRATEGI PEMASARAN DI MALANG GATEWAY

Oleh:
Marcellina Agatha
NIM 21.51.0001

Telah disetujui oleh pembimbing untuk diujikan
Pada tanggal 14 Januari 2026

Menyetujui,
Komisi Pembimbing

Pembimbing I

Dr. Mochamad Husni, S.Pd., MM., MMSI
NIK 01.09.10.036

Pembimbing II

Dr. Weda Adistianaya Dewa, S.Kom., MMSI
NIK 14.02.25.003

Mengetahui,
Ketua Program Studi S-1 Sistem Informasi



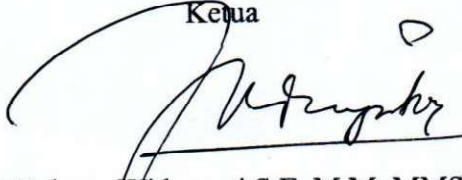
Dr. Dwi Satrio Utsalina, S.Kom., MMSI
NIK 00.12.01.016

LEMBAR PENGESAHAN

ANALISIS DATA WISATAWAN MENGGUNAKAN METODE K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI PEMASARAN DI MALANG GATEWAY

**Skripsi oleh Nama Mahasiswa ini
Telah dipertahankan di depan dewan penguji
Pada tanggal 14 Januari 2026**

Dewan Penguji,
Ketua


Dr. Rahayu Widayanti, S.E., M.M., MMSI
NIK 00.12.01.006

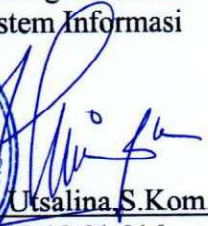
Anggota I


Dr. Dwi Safiroh Utsalina, S.Kom., MMSI
NIK 00.12.01.016

Anggota II


Dr. Mochamad Husni, S.Pd., MM., MMSI
NIK 01.09.10.036

Mengetahui,
Ketua Program Studi
S-1 Sistem Informasi



Dr. Dwi Safiroh Utsalina, S.Kom., MMSI
NIK 00.12.01.016

Mengesahkan,
Ketua STMIK PPKIA
Pradnya Paramita



Dr. Tb. Mohammad Akhriza, S.Si., MMSI
NIK 00.12.01.002

PERNYATAAN ORISINALITAS SKRIPSI

Saya menyatakan dengan sebenar-benarnya bahwa sepanjang pengetahuan saya, di dalam Naskah SKRIPSI ini tidak terdapat karya ilmiah yang pernah diajukan oleh orang lain untuk memperoleh gelar akademik di suatu Perguruan Tinggi, dan tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip dalam naskah ini disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah SKRIPSI ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia SKRIPSI ini digugurkan dan gelar akademik yang saya peroleh (SARJANA) dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku (UU No 20 Tahun 2003, pasal 25 ayat 2 dan pasal 70).

Malang, 14 Januari 2026

Mahasiswa,



Nama : Marcellina Agatha

NIM : 21.51.0001

MOTTO

"Apapun juga yang diperbuat, perbuatlah dengan segenap hati seperti untuk Tuhan, bukan untuk manusia dan yakin bahwa segala perkara dapat ditanggung di dalam Dia yang memberi kekuatan. Jadi, janganlah kuatir akan hari besok, karena hari besok mempunyai kesusahannya sendiri. Kesusahan sehari cukuplah untuk sehari."

ABSTRAK

Agatha, Marcellina. 2025. *Analisis Data Wisatawan Menggunakan Metode K-Means Clustering Untuk Menentukan Strategi Pemasaran Di Malang Gateway*. S1 Sistem Informasi, STMIK PPKIA Pradnya Paramita. Pembimbing: (I) Dr. Mochamad Husni, S.Pdl., MM., MMSI, (II) Dr. Weda Adistianaya Dewa, S.Kom., MMSI.

Kata Kunci: Data Wisatawan, *Clustering*, *K-Means*, *Elbow*, *Silhouette Coefficient*.

Penurunan jangkauan dan kunjungan profil pada akun Instagram @malanggateway disertai penurunan pendapatan dari Rp68 juta menjadi Rp42 juta antara April hingga Agustus 2024 menunjukkan belum maksimalnya strategi promosi digital yang dijalankan karena tidak ada klasterisasi paket yang diminati. Untuk mengatasi hal tersebut, penelitian ini bertujuan mengklasterisasi paket wisata yang paling diminati calon wisatawan dengan menerapkan metode K-Means Clustering. Data yang digunakan berasal dari internal Malang Gateway dan aktivitas sosial media Instagram tahun 2023–2024, dengan total 264 record. Atribut yang dianalisis meliputi tujuan paket wisata, harga, jumlah pendaftar, dan jumlah berminat. Proses analisis melibatkan pre-processing data, normalisasi, serta penentuan jumlah cluster optimal menggunakan metode Elbow dan Silhouette Coefficient. Hasil penelitian menunjukkan bahwa nilai k terbaik adalah 3. Cluster 0 mencakup paket domestik dan internasional harga Rp1,2–5,1 juta dengan jumlah wisatawan tertinggi sekitar 9.000 orang. Cluster 1 terdiri dari paket premium ke Jepang dan Korea seharga Rp11–17 juta dengan wisatawan sekitar 5.000 orang. Sedangkan cluster 2 mencakup paket seperti Bromo, Lombok, dan Bangkok Pattaya dengan harga Rp350 ribu–6 juta dan jumlah wisatawan 5.000–6.000 orang. Hasil clustering ini memberikan landasan strategis dalam menyusun prioritas promosi, serta pemetaan segmen pasar berdasarkan daya tarik dan keterjangkauan harga. Dengan penerapan analisis ini, Malang Gateway dapat menyusun strategi pemasaran yang lebih tepat sasaran.

KATA PENGANTAR

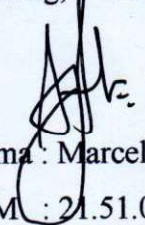
Puji syukur kami panjatkan kehadirat Tuhan Yang Maha Esa atas segala rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi dengan judul **“Analisis Data Wisatawan Menggunakan Metode *K-Means Clustering* Untuk Menentukan Strategi Pemasaran Di Malang Gateway”** dengan baik sebagai syarat untuk memperoleh gelar Sarjana Komputer.

Kami menyadari tanpa adanya dukungan dan kerja sama dari berbagai pihak, kegiatan skripsi ini tidak akan dapat berjalan baik. Untuk itu, kami ingin menyampaikan rasa terima kasih kepada:

1. Tuhan Yang Maha Esa yang telah memberikan petunjuk dan hidayah dalam pembuatan skripsi dan penyusunan laporan sehingga dapat berjalan dengan baik dari awal hingga akhir.
2. Almarhum ayah kami dan ibu kami yang telah memberikan doa dan dukungannya.
3. Kakak dan ipar kami yang telah memberi dorongan dan motivasi.
4. Bapak Dr. Tb. Mohammad Akhriza, S.Si, M.M.S.I, selaku Ketua STMIK PPKIA Pradnya Paramita.
5. Bapak Dr. Dwi Safiroh Utsalina, S.Kom., M.M.S.I, selaku Ketua Program Studi S-1 Sistem Informasi.
6. Bapak Dr. Mochamad Husni, S.PdI., M.M., M.M.S.I dan Bapak Dr. Weda Adistianaya Dewa, S.Kom., MMSI., selaku dosen pembimbing skripsi.
7. Seluruh jajaran Dosen dan Staff STMIK PPKIA Pradnya Paramita Malang dalam membimbing, serta memberikan ilmu yang bermanfaat.
8. Teman-teman seperjuangan dari Program Studi Sistem Informasi dan Teknologi Informasi yang selalu memberikan semangat dan dukungan selama penyelesaian skripsi ini.
9. Dan seluruh pihak yang telah membantu dan mendukung lancarnya pembuatan Skripsi dari awal hingga akhir yang tidak dapat kami sebutkan satu persatu.

Penulis mengharapkan saran dan kritik dari berbagai pihak yang bersifat membangun demi penyempurnaan skripsi ini. Semoga skripsi ini berguna bagi pembaca secara umum dan penulis secara khusus. Akhir kata, penulis ucapkan banyak terima kasih.

Malang, 14 Januari 2026



Nama : Marcellina Agatha

NIM : 21.51.0001

DAFTAR ISI

SKRIPSI.....	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
PERNYATAAN ORISINALITAS SKRIPSI.....	iv
MOTTO	v
ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	6
1.3. Tujuan Penelitian	6
1.4. Batasan Masalah	6
1.5. Kontribusi Penelitian	7
BAB II TINJAUAN PUSTAKA.....	8
2.1. Data <i>Mining</i>	8
2.2. Metode Clustering.....	11
2.3. Algoritma Metode <i>K-Means</i>	12
2.4. Metode <i>Elbow</i>	15
2.5. Python	16
2.6. Google <i>Collaboratory</i> atau Google <i>Colab</i>	16
2.7. Penerapan <i>K-Means</i> untuk <i>Clustering</i> Wisata Terfavorit.....	17
2.8. Evaluasi <i>Clustering</i>	19
BAB III METODOLOGI PENELITIAN	21
3.1. Analisis Permasalahan	21
3.2. Solusi yang Diusulkan	21
3.2.1. Data Mentah.....	22

3.2.2.	Transformasi Data.....	28
3.2.3.	Data Set.....	29
3.2.4.	Pre Processing.....	30
3.2.5.	Data Siap Olah	31
3.2.6.	Pretest nilai k	32
3.2.7.	Tabel Perbandingan $k=2$ s.d. $k=10$ & Kesimpulan Tabel	32
3.2.8.	Grafik Metode <i>Elbow</i> dan Nilai <i>Silhouette Coefficient</i>	33
3.2.9.	Nilai k terbaik.....	33
3.2.10.	<i>Clustering</i> dengan Nilai k yang Diperoleh menggunakan <i>K-Means</i>	34
3.2.11.	Hasil <i>Clustering</i>	34
3.3.	Rancangan Eksperimen.....	35
3.3.1.	Kebutuhan Perangkat	35
3.3.2.	Kebutuhan Data.....	35
BAB IV PENGUJIAN DAN HASIL.....		39
4.1.	Data Mentah.....	39
4.2.	Transformasi Data.....	44
4.3.	Data Set.....	46
4.4.	<i>Pre Processing</i>	47
4.5.	Data Siap Olah	48
4.6.	<i>Pretest</i> nilai k	49
4.7.	Tabel Perbandingan $k=2$ s.d. $k=10$ & Kesimpulan Tabel	51
4.8.	Grafik Metode <i>Elbow</i> dan <i>Silhouette Coefficient</i>	53
4.9.	<i>Clustering</i> dengan Nilai k Optimal menggunakan <i>K-Means</i>	59
4.10.	Hasil <i>Clustering</i>	60
BAB V KESIMPULAN DAN SARAN		68
5.1.	Kesimpulan	68
5.2.	Saran	69
DAFTAR PUSTAKA		

DAFTAR TABEL

No	Nama Tabel	Halaman
3. 1	<i>Sample</i> Data Mentah <i>Guestlist</i> Open Trip Bali.....	23
3. 2	<i>Sample</i> Data Mentah <i>Guestlist</i> Open Trip 3D2N Singapore Malaysia.....	24
3. 3	Data Mentah Analisis <i>K-Means</i> di Malang Gateway tahun 2024.....	26
3. 4	Data Mentah Analisis <i>K-Means</i> di Malang Gateway tahun 2023	27
3. 5	Data Set Analisis <i>K-Means</i> di Malang Gateway	29
3. 6	Data Mentah Analisis <i>K-Means</i> di Malang Gateway tahun 2023	36
3. 7	Data Mentah Analisis <i>K-Means</i> di Malang Gateway tahun 2024.....	37

DAFTAR GAMBAR

No	Nama Gambar	Halaman
1. 1	Jangkauan akun Instagram @malanggateway April 2024	2
1. 2	Kunjungan Profil akun Instagram @malanggateway April 2024	3
1. 3	Jangkauan akun Instagram @malanggateway Agustus 2024	3
1. 4	Kunjungan profil akun Instagram @malanggateway Agustus 2024.....	4
2. 1	KDD.....	9
2. 2	<i>Flowchart</i> Algoritma <i>K-Means</i>	14
3.1.	<i>Flowchart</i> Solusi Yang Diusulkan.....	22
4. 1	<i>Library import Phyton</i>	40
4. 2	<i>Load</i> Data Mentah.....	42
4. 3	<i>Load</i> data Mentah.....	42
4. 4	<i>Display</i> Data Mentah.....	43
4. 5	Proses Transformasi Data.....	45
4. 6	<i>Code</i> Tampilan Data Set.....	46
4. 7	Hasil Output Data Set	47
4. 8	Tahap <i>Pre Processing</i>	47
4. 9	Proses Tampilan <i>Output</i> Data Siap Olah.....	48
4. 10	Hasil Tampilan <i>Output</i> Data Siap Olah.....	49
4. 11	Proses <i>pretest</i> nilai <i>k</i>	50
4. 12	Proses Tampilan Tabel Perbandingan.....	51
4. 13	Tabel Perbandingan Inertia & Silhouette dari <i>k</i>	52
4. 14	Proses Kesimpulan Tabel Perbandingan nilai <i>k</i>	53
4. 15	Kesimpulan Tabel Perbandingan nilai <i>k</i>	53
4. 16	Proses penentuan <i>k</i> optimal dengan <i>Elbow</i> & <i>Silhouette</i>	54
4. 17	<i>Code</i> Visualisasi Grafik <i>Elbow</i>	55
4. 18	Hasil Visualisasi Grafik <i>Elbow</i>	56
4. 19	<i>Code</i> Visualisasi Grafik <i>Silhouette</i>	57
4. 20	Hasil Visualisasi Grafik <i>Silhouette</i>	58
4. 21	<i>Code</i> untuk Kesimpulan Analisis Grafik <i>Elbow</i> dan <i>Silhouette</i>	59
4. 22	Hasil Kesimpulan Analisis Grafik <i>Elbow</i> dan <i>Silhouette</i>	59
4. 23	<i>Code</i> untuk menjalankan <i>K-Means</i> dengan nilai <i>k</i> terbaik	60
4. 24	<i>Code</i> Visualisasi Hasil Clustering dengan Scatter Plot.....	61
4. 25	Scatter Plot Visualisasi Hasil Clustering K-Means.....	62
4. 26	<i>Code</i> Visualisasi Hasil Clustering dengan Tabel.....	63
4. 27	Tabel Visualisasi Hasil Clustering K-Means.....	64
4. 28	<i>Code</i> Visualisasi Hasil Clustering K-Means dengan Grafik	65
4. 29	Grafik Visualisasi Hasil Clustering K-Means	66

DAFTAR LAMPIRAN

No.	Nama Lampiran
1	Riwayat Hidup
2	Berita Acara Bimbingan