

**CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK
MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI
KOTA SURABAYA MENGGUNAKAN ALGORITMA K-
MEANS**

SKRIPSI

Diajukan untuk memenuhi sebagian persyaratan
dalam memperoleh gelar Sarjana Komputer
Program Studi Sistem Informasi



Disusun Oleh:

MUHAMMAD DAFFA

19082010031

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
SURABAYA
2022**

SKRIPSI

**CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK
MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA
SURABAYA MENGGUNAKAN ALGORITMA K-MEANS**

Disusun Oleh:

MUHAMMAD DAFFA
19082010031

**Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi
Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Pada Tanggal 19 Desember 2022**

Pembimbing:

1.



Dhiat Satria Y. K., S.Kom., M.Kom.
NPT. 201198 60 522249

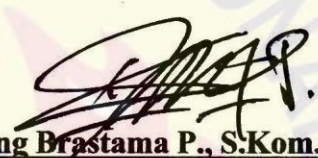
2.



Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 202203 2007

Tim Penguji:

1.



Agung Brastama P., S.Kom., M.Kom.
NIP. 19851124 2021211 003

2.



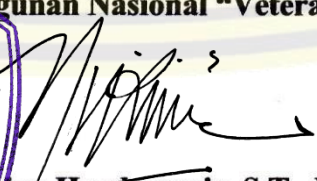
Dr. Eng Agussalim, S.Pd., M.T.
NIP. 19850811 2019031 005

3.



Anindo Saka Fitri, S.Kom., M.Kom.
NPT. 212199 30 325268

Mengetahui,
Dekan Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur



Dr. Novirina Hendrasarie, S.T., M.T.
NIP. 19681126 199403 2 001

LEMBAR PENGESAHAN

**CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK
MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA
SURABAYA MENGGUNAKAN ALGORITMA K-MEANS**

Disusun Oleh:

MUHAMMAD DAFFA
19082010031

**Telah disetujui mengikuti Ujian Negara Lisan Gelombang Desember
Periode 2022 pada Tanggal 19 Desember 2022**

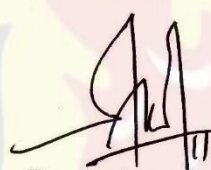
Menyetujui,

Dosen Pembimbing 1



Dhian Satria Y. K., S.Kom., M.Kom.
NPT. 201198 60 522249


Dosen Pembimbing 2



Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 2022032 007

Mengetahui,

**Ketua Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Pembangunan Nasional "Veteran" Jawa Timur**



Nur Cahyo Wibowo, S.Kom., M.Kom.
NIP. 19790317 2021211 002



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
FAKULTAS ILMU KOMPUTER**

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut:

Nama : Muhammad Daffa
NPM : 19082010031
Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 19 Desember 2022 dengan judul:

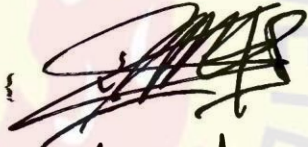
**CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK
MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA
SURABAYA MENGGUNAKAN ALGORITMA K-MEANS**

Oleh karenanya mahasiswa tersebut diatas dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diizinkan untuk membukukan laporan SKRIPSI dengan judul tersebut.

Surabaya, 28 Desember 2022

Dosen penguji yang memeriksa revisi:

1. Agung Brastama P., S.Kom., M.Kom.
NIP. 19851124 2021211 003

{  }

2. Dr. Eng Agussalim, S.Pd., M.T.
NIP. 19850811 2019031 005

{  }

3. Anindo Saka Fitri, S.Kom., M.Kom.
NPT. 212199 30 325268

{  }

Mengetahui,

Dosen Pembimbing 1

Dosen Pembimbing 2


Dhian Satria Y. K., S.Kom., M.Kom.
NPT. 201198 60 522249


Reisa Permatasari, S.T., M.Kom.
NIP. 19920514 2022032 007



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
FAKULTAS ILMU KOMPUTER

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Daffa
NPM : 19082010031
Program Studi : Sistem Informasi

Menyatakan bahwa Judul Skripsi / Tugas Akhir sebagai berikut:

**CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK
MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA
SURABAYA MENGGUNAKAN ALGORITMA K-MEANS**

Bukan merupakan plagiat dari Skripsi / Tugas Akhir / Penelitian orang lain dan juga bukan merupakan Produk / Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Institusi Pendidikan lain. Jika ternyata dikemudian hari pernyataan terbukti benar, maka Saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 19 Desember 2022

Hormat Saya,



Muhammad Daffa
NPM. 1908201xxxx

Judul : CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA SURABAYA MENGGUNAKAN ALGORITMA K-MEANS

Pembimbing 1 : Dhian Satria Yudha Kartika, S.Kom, M.Kom

Pembimbing 2 : Reisa Permatasari, S.T, M.Kom

ABSTRAK

Maraknya tindak kejahatan membuat banyak masyarakat yang resah atau bahkan menjadi korban sehingga mereka membuat laporan pada kepolisian, termasuk pada kejahatan 3C (Curat, Curas, dan Curanmor). Apabila laporan 3C ini bisa ditampilkan ke masyarakat dalam bentuk yang interaktif akan banyak manfaat yang bisa didapat. Penelitian ini bertujuan untuk menampilkan data laporan 3C dalam bentuk tampilan peta interaktif dan dibedakan berdasarkan jenis kejahatan serta klasternya menggunakan algoritma k-means. Dari tampilan peta diharapkan masyarakat dan pihak kepolisian bisa mengambil sikap atas informasi yang dipaparkan. Hasilnya ditemukan bahwa kejahatan Curat, Curas, dan Curanmor telah menyebar di hampir seluruh wilayah Surabaya. Terdapat 2 klaster yang dihasilkan dari algoritma k-means. Hasil data yang diklasterkan hanya dipengaruhi berdasarkan taksiran kerugiannya saja. Dari 79 data hanya 5 data yang berada di klaster 2 yang merupakan data dengan taksiran kerugian yang tinggi sedangkan 74 data terdapat pada klaster 1 yang merupakan klaster dengan taksiran kerugian yang rendah.

Kata kunci:

Kriminalitas, *Web scraping*, Klasterisasi, *K-Means*, *Folium*

KATA PENGANTAR

Segala puji syukur penulis panjatkan atas kehadiran Allah SWT karena berkat limpahan rahmat dan hidayahNya penulis dapat menyelesaikan skripsi yang berjudul “CLUSTERING DATA LAPORAN KRIMINALITAS UNTUK MENGETAHUI PERSEBARAN KASUS KEJAHATAN DI KOTA SURABAYA MENGGUNAKAN ALGORITMA K-MEANS” sebagai salah satu syarat untuk menyelesaikan program studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Skripsi ini bisa terselesaikan juga berkat dorongan dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada pihak yang telah turut serta membantu selama proses pengerjaan skripsi ini. Untuk itu penulis menyampaikan terima kasih kepada:

1. Orang tua, Bapak Totok Suprayitno, S.H dan Ibu Sri Meisaroh Maryanti yang menjadi penyemangat utama serta memberikan doa yang tiada hentinya agar dimudahkan segala urusan saya, termasuk untuk skripsi.
2. Bapak Dhian Satria Yudha Kartika, S.Kom, M.Kom selaku pembimbing I yang telah memberikan arahan dan masukan, terutama untuk hal teknis.
3. Ibu Reisa Permatasari, S.T, M.Kom selaku pembimbing II yang telah memberikan arahan dan masukan, terutama untuk hal penulisan.
4. Pihak Sitipol Polrestabes Surabaya yang memberikan arahan terkait data yang akan diambil dan diolah pada skripsi ini.

5. Bapak Nur Cahyo Wibowo, S.Kom, M.Kom selaku dosen wali yang selalu membantu saya selama kuliah dalam berbagai hal.
6. Seluruh dosen Sistem Informasi yang telah banyak memberikan ilmu yang sangat bermanfaat selama kuliah dan untuk bekal kedepannya.
7. Seluruh teman-teman Sistem Informasi angkatan 2019 yang telah memberi semangat untuk menyelesaikan skripsi ini, terutama untuk kelas A.
8. Rekan-rekan KKN-T Kelompok 2 atas perhatiannya dan senantiasa menyemangati saya dalam menyelesaikan skripsi meskipun terkadang memberikan sedikit hambatan dengan kata-kata “ayo pabji, login login login”.
9. Seluruh pihak yang tidak bisa saya sebutkan satu persatu karena telah memberikan bantuan sekecil apapun.

Semoga Allah SWT membalas kebaikan pihak-pihak diatas. Dalam pengerjaan skripsi ini penulis menyadari masih banyak sekali kekurangan. Semoga skripsi ini bermanfaat bagi semua pihak.

DAFTAR ISI

ABSTRAK	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian	5
1.5 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	8
2.1 Dasar Teori.....	8
2.1.1 Google Maps.....	8
2.1.2 Web Scraping.....	8
2.1.3 Clustering.....	9
2.1.4 K-Means.....	10
2.1.5 Phyton	11
2.1.6 Folium.....	11
2.1.7 Pandas	12
2.1.8 Sklearn	12
2.2 Penelitian Terdahulu	12
BAB III METODOLOGI PENELITIAN.....	16

3.1	Studi Literatur	16
3.2	Analisis Kebutuhan	17
3.2.1	Kebutuhan Data.....	17
3.2.2	Kebutuhan Software dan Hardware	17
3.3	Web Scraping.....	17
3.4	Klasterisasi	18
3.4.1	Persiapan.....	18
3.4.2	Eksplorasi Dataset.....	19
3.4.3	Visualisasi Kemungkinan Adanya Klaster	19
3.4.4	Menentukan Nilai K Optimal.....	19
3.4.5	Validasi Klaster.....	19
3.4.6	Assign Hasil Klaster	20
3.4.7	Visualisasi Hasil Klaster	20
3.5	Visualisasi	20
BAB IV HASIL DAN PEMBAHASAN		22
4.1	Implementasi Kebutuhan	22
4.1.1	Kebutuhan Data.....	22
4.1.2	Kebutuhan Hardware dan Software	22
4.2	<i>Web Scraping</i>	23
4.2.1	Import Data	23
4.2.2	<i>Scrape</i> data Google Maps	23
4.2.3	<i>Split</i> URL	25
4.3	Menghitung Jarak Lokasi Kejadian ke Kantor Polisi	25

4.4	Klasterisasi	27
4.4.1	Persiapan	28
4.4.2	Eksplorasi <i>Dataset</i>	29
4.4.3	Visualisasi Kemungkinan Adanya Klaster.....	29
4.4.4	Menentukan Jumlah K Paling Optimal	30
4.4.5	Validasi Klaster.....	31
4.4.6	<i>Assign</i> Hasil Klaster	32
4.4.7	Visualisasi Hasil Klaster	33
4.5	Visualisasi	33
BAB V PENUTUP.....		40
5.1	Kesimpulan	40
5.2	Saran	41
DAFTAR PUSTAKA.....		42

DAFTAR GAMBAR

Gambar 1. 1 Tampilan Google Maps pada Web Desktop	2
Gambar 3. 1 Alur Metodologi Penelitian.....	16
Gambar 3. 2 Alur Perancangan Model Klasterisasi	18
Gambar 4. 1 Contoh Data yang Akan Diolah	23
Gambar 4. 2 Letak Titik Koordinat Latitude dan Longitude	23
Gambar 4. 3 Source Code untuk scraping data.....	24
Gambar 4. 4 Hasil dari scraping data.....	24
Gambar 4. 5 <i>Source Code Spliting</i>	25
Gambar 4. 6 Hasil dari <i>Spliting</i>	25
Gambar 4. 7 dan Hasil dari Perhitungan Jarak.....	26
Gambar 4. 8 Hasil dari Perhitungan Jarak	26
Gambar 4. 9 <i>Source Code</i> Konversi <i>String</i> ke <i>Float</i>	27
Gambar 4. 10 Hasil Tipe Data	27
Gambar 4. 11 <i>Source Code</i> Perhitungan <i>mean</i> untuk Perlakuan Pertama	27
Gambar 4. 12 Hasil Perhitungan <i>mean</i> untuk Perlakuan Pertama	28
Gambar 4. 13 <i>Source Code</i> Perhitungan <i>mean</i> untuk Perlakuan Kedua.....	28
Gambar 4. 14 Hasil Perhitungan <i>mean</i> untuk Perlakuan Kedua.....	28
Gambar 4. 15 <i>Load Library</i> dan <i>Dataset</i>	29
Gambar 4. 16 Sebagian Data yang Ada pada <i>Dataset</i>	29
Gambar 4. 17 Hasil <i>Plot</i> untuk Visualisasi Kemungkinan Adanya Klaster.....	30
Gambar 4. 18 <i>Source Code</i> untuk Menentukan Jumlah K Paling Optimal	31
Gambar 4. 19 Hasil Visualisasi dari Metode Elbow	31
Gambar 4. 20 Impor <i>Silhouette Coefficient</i>	31
Gambar 4. 21 <i>Source Code</i> dari Perhitungan <i>Silhouette Coefficient</i>	32

Gambar 4. 22 Hasil dari Perhitungan <i>Silhouette Coefficient</i>	32
Gambar 4. 23 Contoh Hasil dari <i>Assign</i> Klaster	33
Gambar 4. 24 Grafik Visualisasi Hasil Klaster.....	33
Gambar 4. 25 Impor <i>Library</i> dan Data pada Proses Visualisasi	34
Gambar 4. 26 <i>Source Code</i> untuk <i>mapping</i> Variabel <i>kd_warna</i>	34
Gambar 4. 27 Tampilan Kota Surabaya pada Folium.....	35
Gambar 4. 28 <i>Source Code</i> Visualisasi.....	35
Gambar 4. 29 Hasil Visualisasi.....	36
Gambar 4. 30 <i>Source Code</i> untuk Memisah Tiap Jenis Kejahatannya	36
Gambar 4. 31 <i>Source Code</i> Visualisasi CURAT	37
Gambar 4. 32 Hasil Visualisasi CURAT	37
Gambar 4. 33 <i>Source Code</i> Visualisasi CURAS	37
Gambar 4. 34 Hasil Visualisasi CURAS	38
Gambar 4. 35 <i>Source Code</i> Visualisasi CURANMOR.....	38
Gambar 4. 36 Hasil Visualisasi CURANMOR.....	38

DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu	12
---------------------------------------	----