

---

## ***CLUSTERING RUMAH ISOLASI DI KOTA SURAKARTA***

### ***CLUSTERING HOUSE INSULATION CITY OF SURAKARTA***

Ardymulya Iswardani<sup>1</sup>, Tri Rochmadi<sup>2</sup>, Nugroho Arif Sudiby<sup>3</sup>,  
Tiyang Ganang Wicaksana<sup>4</sup>, Chaerina Barabas<sup>5</sup>

<sup>1,3,4</sup>Universitas Duta Bangsa, Surakarta

<sup>2,5</sup>Universitas Alma Ata, Yogyakarta

e-mail: <sup>1</sup>ardymulya@udb.com, <sup>2</sup>trirochmadi@almaata.ac.id, <sup>3</sup>nugrohoarifs@udb.ac.id,  
<sup>4</sup>tiyang\_ganang@udb.ac.id, <sup>5</sup>203100156@almaata.ac.id

#### **ABSTRAK**

Penelitian ini bertujuan untuk melakukan *clustering* rumah isolasi bagi warga kota Surakarta yang melaksanakan isolasi mandiri karena infeksi Covid-19. Rumah isolasi tersebut diproses dengan algoritma *k-means* untuk analisis berikutnya. Rentang waktu pengambilan data 1 hari periode PPKM darurat. Data tersebut di *cluster* dengan bahasa pemrograman R yang bersifat *open source*. Algoritma *k-means* mengevaluasi jarak antar data berdasarkan tingkat kemiripan dengan *centroid*. Algoritma yang bekerja secara mandiri ini menghasilkan visualisasi data *clustering* rumah isolasi warga kota Surakarta yang melakukan isolasi mandiri karena terinfeksi Covid-19 dengan nilai optimal *cluster* pada 54 kelurahan yaitu c1 sebanyak 2 kelurahan, c2 = 3 kelurahan, c3 = 8 kelurahan dan c4 = 41 kelurahan. Hasil *clustering* berdasarkan data tersebut menunjukkan bahwa kota Surakarta memiliki 2 (3%) zona merah, 3 (5%) zona oranye dan 8 (14%) zona kuning dan 41 (75%) zona hijau. Hal tersebut menunjukkan bahwa digunakannya rumah warga yang dialihfungsikan sebagai fasilitas isolasi mandiri dan upaya pembatasan kegiatan masyarakat dapat mengendalikan penyebaran Covid-19 di kota Surakarta.

**Kata Kunci:** *data mining, clustering, k-means, Covid, zonasi*

#### **ABSTRACT**

This study aims to conduct clustering of isolation houses for residents of the city of Surakarta who carry out self-isolation due to Covid-19 infection. The isolation house was processed with the k-means algorithm for subsequent analysis. The data collection time span is 1 day for the emergency PPKM period. The data is clustered with the R programming language which is open source. The k-means algorithm evaluates the distance between data based on the degree of similarity to the centroid. This independently working algorithm produces data visualization of clustering isolation houses for residents of the city of Surakarta who are self-isolating because they are infected with Covid-19, namely the optimal value of clusters in 54 villages, namely c1 as many as 2 villages, c2 = 3 villages, c3 = 8 villages and c4 = 41. Ward. The results of clustering based on these data show that the city of Surakarta has 2 (3%) red zones, 3 (5%) orange zones and 8 (14%) yellow zones and 41 (75%) green zones. This shows that the use of residents' houses which have been converted as self-isolation facilities and efforts to limit community activities can control the spread of Covid-19 in the city of Surakarta.

**Keywords:** *data mining, clustering, k-means, covid, zone*

## 1.PENDAHULUAN

Deklarasi Kejadian Luar Biasa (KLB) pandemi Covid-19 di kota Surakarta pada hari Jumat tanggal 13 Maret 2020 oleh Wali kota Surakarta F.X Hadi Rudyatmo merupakan bentuk status yang menyatakan suatu penyakit yang dapat merebak menjadi wabah. Penyebutan KLB pada dasarnya merupakan peringatan dini sebelum peristiwa tersebut menjadi wabah di tengah masyarakat.

Covid-19 yang dinyatakan sebagai KLB di kota Surakarta memantik berbagai elemen masyarakat untuk bahu membahu menangani pengendalian Covid-19. Salah satu upaya pengendalian penyebaran Covid-19 yaitu dengan menindaklanjuti instruksi Gubernur Jawa Tengah Nomor 1 Tahun 2020 tentang Pemberdayaan Masyarakat dalam Percepatan Penanganan Covid-19 di tingkat Rukun Warga (RW) melalui Pembentukan “Satgas Jogo Tonggo”[1] dengan melaksanakan bimbingan teknis aplikasi Jogo Tonggo[2]. Bimbingan teknis ini bertujuan sebagai media pelaporan monitoring dan evaluasi kasus Covid-19 di wilayah Provinsi Jawa Tengah.

Aplikasi Jogo Tonggo hingga saat ini telah mengalami banyak pengembangan yaitu berbasis web, mobile dan dukungan *dashboard* sebagai pusat data Covid 19 sesuai wilayah, fitur yang dulu hanya 4 (empat) bidang seperti kesehatan, ekonomi, hiburan dan sosial dan keamanan telah berkembang dengan adanya fitur PPKM Mikro dan Pendataan Pemudik. Kedua fitur tambahan tersebut tersedia di awal tahun 2021.

Pengembangan tersebut tidak lepas dari perhatian pemerintah provinsi tentang pentingnya Informasi yang berkualitas dari banyaknya data (*piles of data*) yang dilaporkan baik melalui sistem maupun bentuk laporan manual lainnya. Data yang terkumpul tersebut menjadi bongkahan data yang menyimpan pengetahuan untuk dukungan pengambilan keputusan dan manfaat pengolahan bongkahan data yaitu untuk proses pembelajaran dalam mengenali pola yang terjadi sebagai dasar untuk mengenali pola baru yang dapat muncul di masa mendatang.

Ada berbagai jenis teknik pengolahan data salah satunya yaitu *data mining* (menggali data). *Data mining* merupakan suatu teknik pengolahan data dengan memanfaatkan bongkahan data yang berasal dari waktu lampau untuk kemudian diproses sedemikian agar mengungkap pengetahuan dari bongkahan data tersebut, istilah untuk menyebut rangkaian proses tersebut dikenal sebagai *knowledge discovery* [3]

Tahapan dalam melakukan penggalian pengetahuan dalam *data mining* yaitu: 1.*data cleaning and integration*, 2.*data selection and transformation*, 3.*data mining* dan 4.*evaluation and presentation* [4].

*Data mining* sendiri merupakan teknik yang secara khusus digunakan untuk menangani data dengan jumlah besar. Teknik *data mining* yang digunakan dalam penelitian ini yaitu *clustering* dengan algoritma *k-means* [5].

Algoritma ini berupaya untuk menghasilkan *cluster* dari kelompok data yang dijadikan input algoritma *k-means*. Algoritma *k-means* berjenis algoritma tidak terawasi (*unsupervised learning*) yang bertujuan untuk mempartisi suatu objek data berdasarkan tingkat kemiripan (*similarity*) dengan centroidnya. Suatu objek data dengan kesamaan tertentu akan dikelompokkan dalam satu *cluster* yang sama dan sebaliknya apabila suatu objek data tersebut berbeda akan dikelompokkan dalam *cluster* yang lain [6].

Data rumah isolasi mandiri yang tersimpan pada aplikasi Jogo Tonggo diperbaharui secara harian oleh setiap operator aplikasi Jogo Tonggo pada tingkat RW yang kemudian data tersebut di verifikasi oleh petugas kelurahan masing-masing. Data rumah isolasi mandiri pada aplikasi Jogo Tonggo tersebut juga disinkronisasikan dengan Sistem Informasi Tata Ruang dan Infrastruktur Perencanaan Kota Surakarta (INTIP) pada alamat: <https://intip.surakarta.go.id> pada menu Intip Covid di alamat <https://covid.intip.surakarta.go.id>.

Data yang digunakan dalam penelitian ini berasal dari aplikasi GIS (*Geografis Information System*) di alamat <https://covid.intip.surakarta.go.id> yang mengandung visualisasi peta wilayah kelurahan berdasarkan zonasi resiko Covid 19 (Tabel 1), selain visualisasi peta tersebut juga menampilkan data legenda di masing-masing kelurahan.

Penelitian ini memanfaatkan data dari GIS tersebut karena sesuai dengan kriteria zonasi resiko Covid yang telah ditetapkan oleh Badan Nasional Penanggulangan Bencana (BNPB) [7].

BNPB mengelompokkan zonasi wilayah berdasarkan resiko Covid-19 menjadi 4 (empat) level yaitu: wilayah dengan resiko Covid dengan kriteria zona hijau (*H*), zona kuning(*K*), zona oranye(*O*) dan zona merah(*M*).

Penelitian ini bertujuan untuk mengungkap pengetahuan dengan melakukan analisis *cluster* kelompok rumah yang digunakan sebagai isolasi mandiri oleh warga yang terkonfirmasi positif Covid-19 dengan penggunaan teknik *data mining* serta bertujuan untuk memperkaya pendidikan studi islam di era perkembangan teknologi yang semakin maju dengan demikian umat islam dapat tetap beribadah dan berkarya sesuai dengan syariah islam meskipun berbantuan teknologi.

## 2.LITERATURE REVIEW

*Clustering* banyak digunakan di berbagai bidang untuk mengatasi berbagai permasalahan diantaranya pada penelitian atas produksi kacang tanah nasional [8]. Penelitian tersebut menggunakan *clustering* untuk mengelompokkan wilayah penghasil kacang tanah dengan kriteria tinggi dan rendah. Tujuan melakukan analisis *cluster* pada produksi kacang tanah tersebut bermanfaat untuk mengetahui kapasitas kemampuan wilayah dalam memproduksi kacang tanah untuk suplai kacang tanah di seluruh wilayah Indonesia. Indikator yang digunakan sebagai input algoritma *clustering* pada penelitian tersebut yaitu luas lahan (*Ha*) dan jumlah produksi (*ton*). Penelitian tersebut menunjukkan bahwa produksi kacang tanah nasional masih didominasi di wilayah Pulau Jawa yaitu Jawa Barat, Jawa Tengah dan Jawa Timur.

Pemanfaatan lain dari *clustering* juga dilakukan untuk keperluan peramalan (*forecast*). Peramalan yang dilakukan oleh penelitian ini diterapkan untuk meramalkan jumlah penduduk miskin di Indonesia akibat pandemi Covid-19 [9]. Penelitian ini menggabungkan algoritma *clustering* untuk mempartisi data penduduk miskin dan menggunakan metode *Double Exponential Smoothing* (DES) untuk melakukan peramalan. Pemodelan data *time series*, memanfaatkan data jumlah penduduk miskin dari tahun 2006 sampai dengan 2020. Kombinasi kedua teknik pengolahan data tersebut menghasilkan suatu peramalan bahwa akibat pandemi Covid-19 akan meningkatkan jumlah penduduk miskin setiap tahunnya di Indonesia secara khusus dan dunia secara umum.

Kemudahan dan kemampuan algoritma *clustering* banyak digunakan tidak hanya pada kedua kasus penelitian diatas namun juga dapat diterapkan pada analisis serangan *cyber* yang memanfaatkan data lalu lintas jaringan. Penelitian dengan genre simulasi serangan *Denial Of Service* (DoS) bertujuan untuk mengelompokkan tingkat bahaya sedang dan tingkat bahaya tinggi pada atribut *length* dan *tcp length* [10].

Ketiga penelitian tersebut menunjukkan penggunaan metode *clustering* bermanfaat sebagai analisis awal dan penggunaan metode lain bertujuan untuk meningkatkan kualitas hasil pengolahan data pada bidang *machine learning*, *pattern recognition* dan lain sebagainya.

*Clustering* merupakan teknik untuk membagi data dalam suatu himpunan data ke dalam kelompok data yang memiliki kesamaan [11]. Istilah *clustering* digunakan dalam menyebut proses pengelompokkan data, sedangkan teknik untuk mengelompokkan data yaitu algoritma *k-means*.

*k-means* secara umum menggunakan 5 (lima) tahapan proses yang diuraikan sebagai berikut:

1. *Inisialisasi*: Tentukan nilai  $K$  sebagai jumlah *cluster* yang diinginkan.
2. Pilih  $K$  data dari *dataset* sebagai *centroid*
3. Alokasikan semua data ke *centroid* terdekat dengan metrik jarak yang telah ditetapkan.
4. Hitung kembali *centroid*  $C$  berdasarkan data yang mengikuti masing-masing *cluster*.
5. Ulangi langkah 3 dan 4 hingga kondisi *konvergen* tercapai (tidak ada data yang berpindah).

Algoritma *k-means* merupakan jenis metode *clustering* non hirarki yang bekerja dengan mengobservasi suatu objek data dengan *centroid* berdasarkan jarak [12].

Pengukuran jarak data yang umum digunakan yaitu jarak *Euclidean*[13]. (Persamaan 1). Dimana:  $D$  adalah jarak antara data  $X$  dan data  $Y$ , dan  $|\cdot|$  adalah nilai absolut.  $N$  adalah jumlah fitur (dimensi) data.

$$D(x, y) = \|x - y\|_2 = \sqrt{\sum_{j=1}^N |x_j - y_j|^2} \quad (1)$$

Gambaran umum *clustering*, *k-means* dan jarak *euclidean* merupakan tiga hal utama perhatian pengguna metode. Kemudahan dan kecepatan penerapan metode juga merupakan pertimbangan dalam menentukan metode dalam penelitian ini.

Metode *elbow* merupakan metode untuk menentukan jumlah *cluster* perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik [14]

Zonasi resiko Covid adalah kriteria tingkat zonasi daerah berdasarkan warna yang dilihat dari transmisi atau penyebarannya. Gugus Tugas Nasional Penanganan Covid-19 membagi zonasi resiko covid menjadi 4 (empat) level yaitu:

1. Level 1 (Zona Hijau) apabila resiko penyebaran virus ada, akan tetapi tidak ada kasus positif.
2. Level 2 (Zona Kuning) apabila penyebaran Covid-19 dapat terkendali dan tetap ada kemungkinan transmisi.
3. Level 3 (Zona Oranye) apabila resiko penyebaran Covid-19 pada level ini tinggi dan potensi virus tidak terkendali.
4. Level 4 (Zona Merah) apabila penyebaran virus SARS-CoV-2 atau Corona jenis baru penyebab Covid-19 tidak terkendali. Transmisi lokal sudah terjadi dengan cepat, wabah menyebar secara luas dan banyak kluster-kluster baru.

Pasien Covid-19 adalah sumber terjadinya infeksi bagi orang yang sehat. Pasien ini dikenal sebagai pembawa (*carrier*). Pasien yang terinfeksi Covid-19 umumnya disebabkan oleh *droplet* (percikan cairan) karena batuk, bersin, meludah dimana cairan tersebut melayang di udara untuk rentang waktu tertentu [15].

Langkah mitigasi Covid-19 banyak dilakukan sebagai upaya pengendalian penyebaran Covid-19 diantaranya yaitu mitigasi berbasis komunitas dan *timing* [16].

Pasien yang memiliki atau ditetapkan menjadi terkonfirmasi ditangani dengan berbagai cara dari medis hingga mitaigasi berbasis komunal. Mitigasi komunal ialah upaya melokalisir penyebaran, pengendalian Covid-19 ke komunitas lain.

Isolasi atau karantina ialah upaya yang dilakukan kelompok atau individu yang sehat dan atau yang sakit agar tidak terinfeksi dan atau menginfeksi orang lain sebagai upaya deteksi dini dan atau mengurangi resiko terjadinya penularan dalam skala tertentu [17].

### 3.METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini melibatkan 3 (tiga) tahap (gambar 1) yang dijelaskan sebagai berikut ini:



Gambar 1. Metode Penelitian

#### 1) Pra Proses (*Input*)

Merupakan proses awal *clustering* yang melibatkan proses:

##### - Pengumpulan data

Proses pengumpulan data rumah yang dijadikan isolasi mandiri oleh warga kota Surakarta yang dikumpulkan dan dimasukkan dalam format CVS (*comma separated values*) dari <https://covid.intip.surakarta.go.id> yang mengandung visualisasi geografis Wilayah Kelurahan Kota Surakarta yang dilakukan pada hari Sabtu, 17 Juli 2021

##### - *Data cleaning*

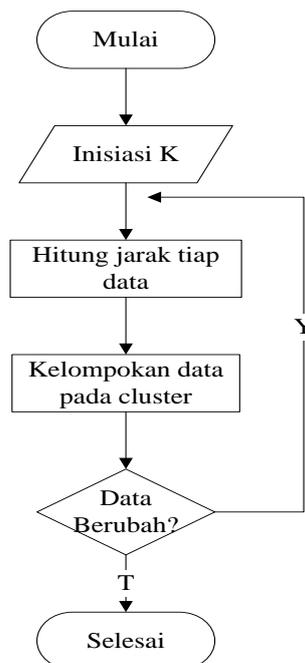
Proses membersihkan data karakter input (bila ada). Hal ini perlu dilakukan agar data yang akan dimasukkan dapat diproses dengan baik.

##### - *Data selection*

Proses pemilihan kriteria yang dijadikan input algoritma *clustering*. Hal ini dilakukan karena hanya data yang dibutuhkan adalah data zonasi resiko Covid pada tingkat RT.

#### 2) Proses Data Mining (*Process*)

Proses data mining merupakan tahap utama dari rangkaian penelitian. Tahap ini data zonasi resiko Covid-19 Kota Surakarta dikelompokkan berdasarkan *centroid* yang ditentukan sebanyak  $k=4$  *cluster*. Proses *clustering* dengan algoritma *k-means* menggunakan alur proses (gambar 2).



Gambar 2. Proses *Clustering*

TABEL I  
KELURAHAN KOTA SURAKARTA

KEL	H	K	O	M	KEL	H	K	O	M
BALUWARTI	21	17	0	0	KRATONAN	29	6	0	0
BANJARSARI	72	31	0	0	LAWEYAN	2	7	1	0
BANYUANYAR	17	25	7	0	MANAHAN	60	2	0	0
BUMI	21	7	0	0	MANGKUBUMEN	42	16	0	0
DANUKUSUMAN	43	12	3	0	MOJO	47	11	0	0
GAJAHAN	21	9	2	0	MOJOSONGO	135	53	4	1
GANDEKAN	13	20	3	0	NUSUKAN	84	57	2	0
GILINGAN	108	4	0	0	PAJANG	36	41	11	0
JAGALAN	30	22	11	0	PANULARAN	32	16	0	0
JAJAR	44	1	0	0	PASARKLIWON	23	13	0	0
JAYENGAN	21	9	0	0	PENUMPING	21	7	0	0
JEBRES	55	72	1	0	PUCANGSAWIT	39	14	3	0
JOGLO	32	24	4	0	PUNGGAWAN	20	9		0
JOYOSURAN	53	2	0	0	PURWODININGRATAN	24	10	1	0
JOYOTAKAN	14	18	1	0	PURWOSARI	35	15	1	0
KADIPIRO	68	28	11	0	SANGKRAH	41	17	0	0
KAMPUNGBARU	14	7	0	0	SEMANGGI	73	14	0	0
KARANGASEM	28	8	0	0	SERENGAN	35	28	4	0
KAUMAN	16	5	0	0	SETABELAN	14	15	2	0
KEDUNGLUMBU	12	12	6	0	SEWU	17	16	3	0
KEMLAYAN	19	4	0	1	SONDAKAN	24	18	8	1
KEPATIHAN KULON	13	6	1	0	SRIWEDARI	14	10	1	0
KEPATIHAN WETAN	13	5	0	0	SUDIROPRAJAN	32	3	0	0
KEPRABON	17	7	2	0	SUMBER	56	18	2	0
KERTEN	19	24	5	0	TEGALHARJO	20	8	4	0
KESTALAN	11	9	0	0	TIMURAN	15	7	0	0
KETELAN	20	7	0	0	TIPES	58	9	2	0

Proses dimulai dengan menentukan  $k = 4$ , yang kemudian setiap dihitung jaraknya untuk menentukan derajat kesamaan (*similarity*). Data yang memiliki kesamaan dimasukkan dalam suatu cluster sedangkan data yang berbeda dimasukkan dalam *cluster* lainnya hingga seluruh data tidak berubah.

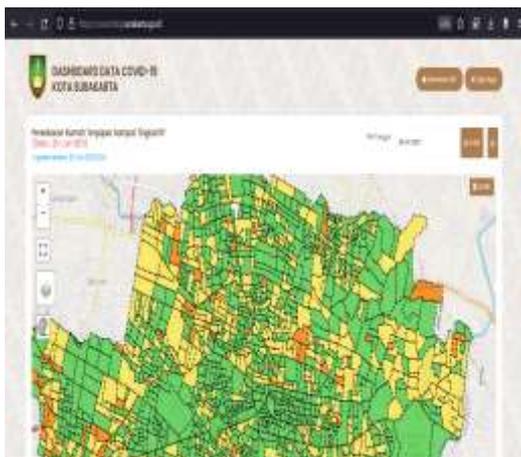
### 3) Analisis Cluster (Output)

Proses evaluasi dan menyajikan laporan hasil *clustering* yang telah dilakukan. Tahap ini menjelaskan representasi data berdasarkan metode *clustering* yang dipilih dalam penelitian deskriptif kualitatif.

## 4. PEMBAHASAN MASALAH

### 1. Pra Proses (Input)

Pengumpulan data sebagai bahan *clustering* yaitu data zonasi resiko covid-19 di wilayah kelurahan kota Surakarta yang diambil pada Hari Sabtu, Tanggal 17 Juli 2021 yang berisi data RT Hijau, RT Kuning, RT Oranye dan RT Merah dari <https://covid.intip.surakarta.go.id> (gambar 3). Data yang telah dikumpulkan kemudian diproses sedemikian hingga menjadi data dalam format yang dibutuhkan lihat (tabel 1). Data pada tabel 1 tersebut merupakan data yang telah mengalami proses pembersihan data dan pemilihan kriteria *clustering* yaitu RT Hijau, RT Kuning, RT Oranye dan RT Merah.



Gambar 3. Dashboard Covid

## 2. Proses Data Mining (*Process*)

Proses *clustering* dalam penelitian ini menggunakan bahasa pemrograman R. Tahapan proses *clustering* dengan bahasa pemrograman R dimulai dengan menyiapkan *library* yang mendukung proses *clustering*

### - Instalasi *library*

*Library* adalah koleksi fungsi yang diperlukan dalam melakukan suatu proses (gambar 4). Penelitian ini bertujuan untuk melakukan *clustering* dengan demikian koleksi yang dibutuhkan yaitu 4 (empat) koleksi dibawah ini:

```
library(tidyverse)
library(cluster)
library(factoextra)
library(gridExtra)
```

Gambar 4. *Library*

keempat perintah diatas akan menjalankan proses instalasi *library* langsung ke *repository* R melalui internet.

### - Load *library*

Suatu mekanisme yang disediakan bahasa pemrograman R agar memudahkan pengguna yang secara intensif menggunakan *library* tersebut dapat langsung menggunakan dengan memuat *library* tersebut pada memori (gambar 5).

```
install.packages("tidyverse")
install.packages("cluster")
install.packages("factoextra")
install.packages("gridExtra")
```

Gambar 5. Load *library*

keempat ialah instruksi yang dijalankan setelah proses instalasi *library* selesai dan kemudian memuat paket tersebut untuk digunakan selama proses *clustering*. perintah diatas

### - Proses *clustering*

(Gambar 6) Perintah `read.csv` berfungsi me-load data pada memory bahasa R yang kemudian data tersebut dibersihkan dengan fungsi `na.omit`. Data yang telah dibersihkan selanjutnya dibuat skala agar data tersebut normal dengan perintah `scale`. Data yang telah diskala

kemudian di-cluster dengan fungsi *k-means*. Proses *clustering* yang digunakan pada penelitian ini menggunakan  $k = 4$  sebagai *centroid* dan inisiasi awal menggunakan nilai yang direkomendasikan secara umum yaitu 25.

```

> dataPPKM =
  read.csv("170721.csv",
    row.names=1,
    header=TRUE)
> dataPPKM <-
  na.omit(dataPPKM)
> dataPPKM <-
  scale(dataPPKM)
> k4 <- kmeans(dataPPKM,
  centers=4, nstart=25)
> fviz_cluster(k4,
  data=dataPPKM)
  
```

Gambar 6. Skrip K-Means

Skrip bahasa pemrograman R tersebut diatas menjelaskan tahapan pencarian pengetahuan dalam proses data mining seperti *data cleaning*, *data selection*, *data mining* dan *data evaluation*. Tahap data evaluation ialah tahap analisis yang diuraikan pada point analisis *cluster*. Data yang akan diproses oleh algoritma *clustering* dinormalisasi terlebih dulu agar meminimalkan jarak/scale (tabel II). Tahap ini bertujuan agar kualitas *cluster* yang dihasilkan baik.

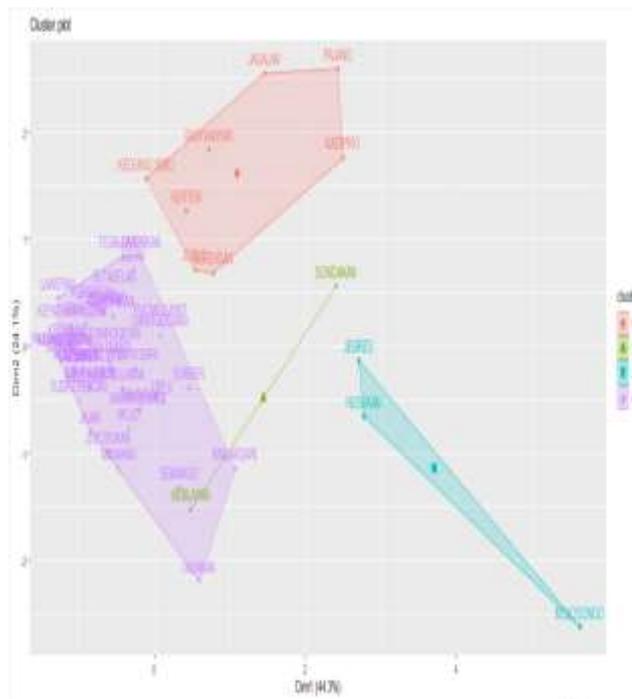
TABEL II  
SKALA JARAK

	H	K	O
BALUWARTI	-0.476	1.107	-0.715
BANJARSARI	1.918	3.460	-0.715
BUMI	-0.476	-0.574	-0.715
DANUKUSUMAN	0.556	0.266	1.872
GAJAHAN	-0.476	-0.238	1.009
GANDEKAN	-0.852	1.611	1.872
GILINGAN	3.607	-1.078	-0.715
JAJAR	0.603	-1.582	-0.715
JAYENGAN	-0.476	-0.238	-0.715
JOYOSURAN	1.026	-1.414	-0.715
JOYOTAKAN	-0.805	1.275	0.147
KAMPUNGBARU	-0.805	-0.574	-0.715
KARANGASEM	-0.148	-0.406	-0.715
KAUMAN	-0.711	-0.910	-0.715
KEPATIHAN KULON	-0.852	-0.742	0.147
KEPATIHAN WETAN	-0.852	-0.910	-0.715
KEPRABON	-0.664	-0.574	1.009
KESTALAN	-0.946	-0.238	-0.715
KETELAN	-0.523	-0.574	-0.715
KRATONAN	-0.101	-0.742	-0.715
LAWEYAN	-1.368	-0.574	0.147

	H	K	O
MANAHAN	1.354	-1.414	-0.715
MANGKUBUMEN	0.509	0.939	-0.715
MOJO	0.744	0.098	-0.715
PANULARAN	0.040	0.939	-0.715
PASARKLIWON	-0.382	0.435	-0.715
PENUMPING	-0.476	-0.574	-0.715
PUCANGSAWIT	0.369	0.603	1.872
PUNGGAWAN	-0.523	-0.238	1.009
PURWODININGRATAN	-0.335	-0.070	0.147
PURWOSARI	0.181	0.771	0.147
SANGKRAH	0.463	1.107	-0.715
SEMANGGI	1.965	0.603	-0.715
SETABELAN	-0.805	0.771	1.009
SEWU	-0.664	0.939	1.872
SRIWEDARI	-0.805	-0.070	0.147
SUDIROPRAJAN	0.040	-1.246	-0.715
SUMBER	1.167	1.275	1.009
TEGALHARJO	-0.523	-0.406	2.734
TIMURAN	-0.758	-0.574	-0.715
TIPES	1.260	-0.238	1.009

Hasil penggunaan fungsi *scale*, jarak antar data menjadi serupa satu sama lain.



Gambar 7. K=4

Berdasarkan hasil *clustering* setiap RT pada masing-masing kelurahan (tabel III) anggota pada setiap *cluster* sebagai berikut:

TABEL III  
ANGGOTA CLUSTER

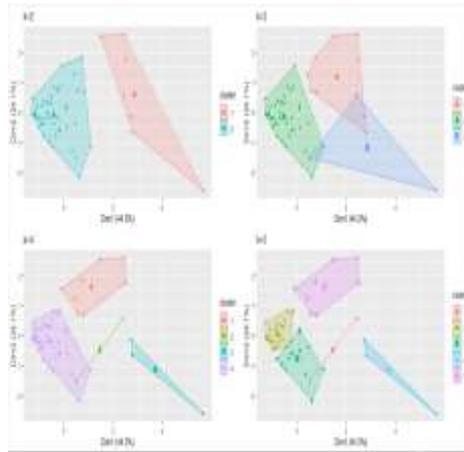
C	KEL	H	K	O
C1	KEMLAYAN	19	4	0
C1	SONDAKAN	24	18	8
C2	MOJOSONGO	135	53	4
C2	JEBRES	55	72	1
C2	NUSUKAN	84	57	2
C3	KEDUNG LUMBU	12	12	6
C3	JOGLO	32	24	4
C3	SERENGAN	35	28	4
C3	KERTEN	19	24	5
C3	BANYUANYAR	17	25	7
C3	JAGALAN	30	22	11
C3	PAJANG	36	41	11
C3	KADAPIRO	68	28	11
C4	BALUWARTI	21	17	0
C4	BANJARSARI	72	31	0
C4	BUMI	21	7	0
C4	DANUKUSUMAN	43	12	3
C4	GAJAHAN	21	9	2
C4	GANDEKAN	13	20	3
C4	GILINGAN	108	4	0
C4	JAJAR	44	1	0
C4	JAYANGAN	21	9	0
C4	JOYOSURAN	53	2	0
C4	JOYOTAKAN	14	18	1
C4	KAMPUNGBARU	14	7	0
C4	KARANGASEM	28	8	0
C4	KAUMAN	16	5	0
C4	KEPATIHAN WETAN	13	5	0
C4	KEPRABON	17	7	2
C4	KESTALAN	11	9	0
C4	KETELAN	20	7	0
C4	KRATONAN	29	6	0
C4	LAWEYAN	2	7	1
C4	MANAHAN	60	2	0
C4	MANGKUBUMEN	42	16	0
C4	MOJO	47	11	0
C4	PANULARAN	32	16	0
C4	PASARKLIWON	23	13	0
C4	PENUMPING	21	7	0
C4	PUCANGSAWIT	39	14	3
C4	PUNGGAWAN	20	9	2
C4	PURWODININGRATAN	24	10	1
C4	PURWOSARI	35	15	1
C4	SANGKRAH	41	17	0
C4	SEMANGGI	73	14	0
C4	SETABELAN	14	15	2
C4	SEWU	17	16	3
C4	SRIWEDARI	14	10	1
C4	SUDIROPRAJAN	32	3	0
C4	SUMBER	56	18	2
C4	TEGALHARJO	20	8	4
C4	TIMURAN	15	7	0
C4	TIPES	58	9	2
C4	KEPATIHAN KULON	13	6	1

Hasil proses *clustering* divisualisasikan (gambar 7).

### 3. Analisis Cluster (*Output*)

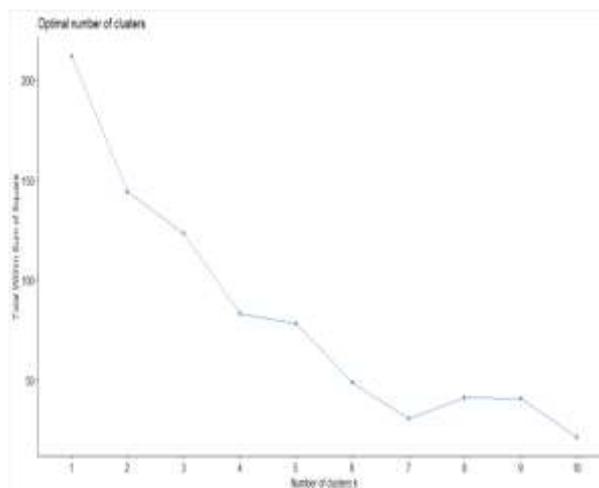
Analisis cluster dilakukan dengan 2 (dua) bagian, yaitu:

- Perbandingan  $k$   
Perbandingan centroid bertujuan untuk mengetahui jumlah  $k$  dengan nilai  $k=2$ ,  $k=3$ ,  $k=4$  dan  $k=5$ . (gambar 8)



Gambar 8. Perbandingan K

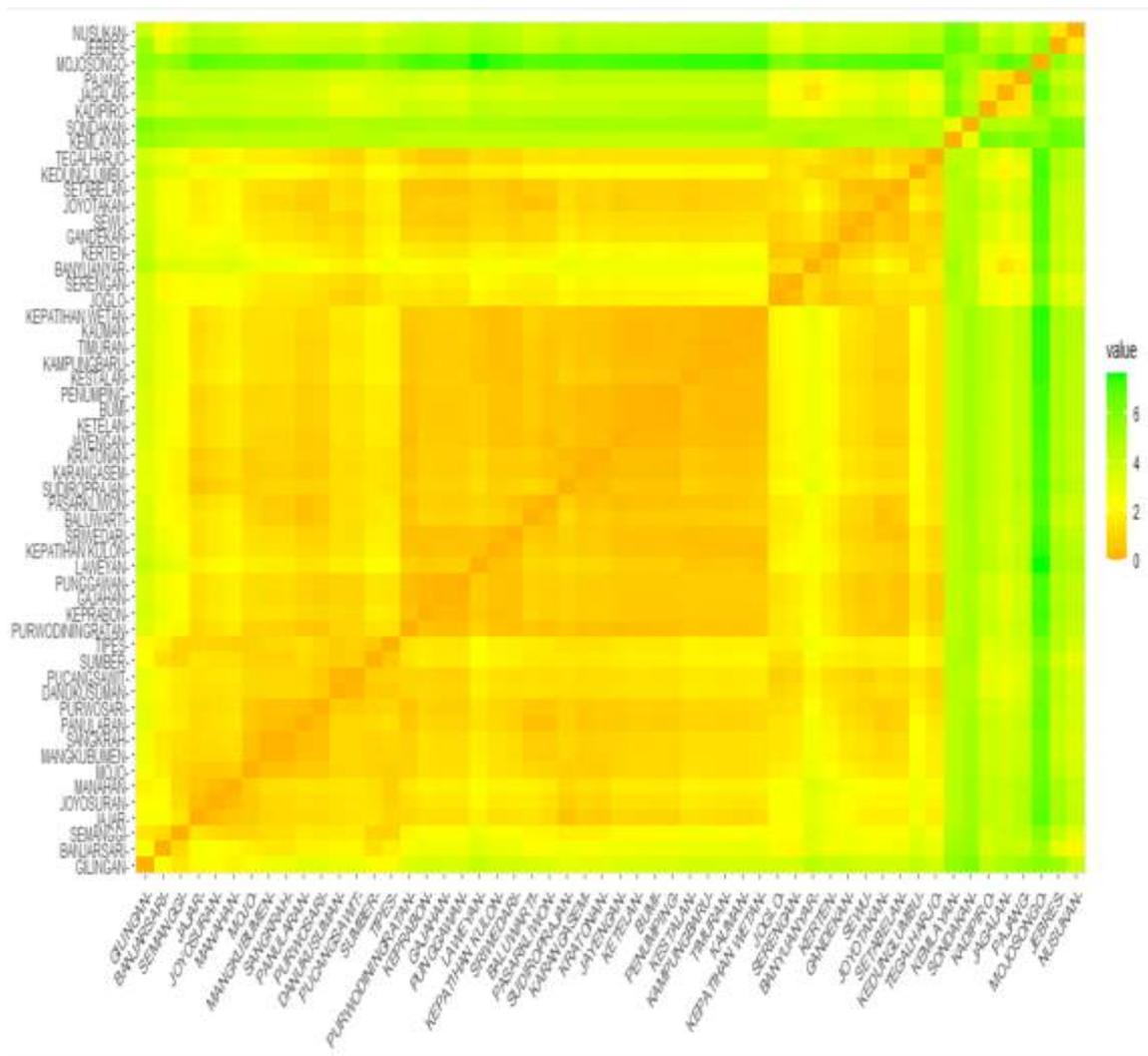
Fungsi membandingkan nilai K tersebut bertujuan untuk mengetahui jarak optimal antar data yang tergabung dalam suatu *cluster*. Hasil perbandingan *cluster* menunjukkan  $k = 4$  merupakan nilai *cluster* optimal. Jumlah  $k = 4$  tersebut merupakan hasil analisis dengan menggunakan metode Elbow (gambar 9).



Gambar 9. Elbow

Gambar tersebut menunjukkan sudut yang membentuk siku (*elbow*) ditunjukkan dengan nilai= 4. Nilai ini menunjukkan jumlah *cluster* optimal  $k= 4$ .

Visualisasi jarak proses *clustering* dibagi menjadi tiga bagian yaitu rendah, sedang dan tinggi (gambar 10).



Gambar 10. Warna Jarak

Gambar tersebut menunjukkan visualisasi jarak antar data pada *cluster*.

- Analisis Cluster

Visualisasi jarak dikelompokkan menjadi kategori warna:

- Rendah warna merah karena zona merahnya sedikit (justru ini yang menjadi pusat perhatian)
- Sedang warna kuning karena zona kuning & oranya dikategorikan sebagai centroid sedang.
- Tinggi warna hijau karena zona hijau yang menjadi pusat perhatian berikutnya.

C1 :

Kedua wilayah ini merupakan pusat bisnis penopang perekonomian kota Surakarta yang menghasilkan batik. Fungsinya sebagai pusat bisnis menunjukkan adanya kegiatan aktifitas perekonomian yang padat yang berpotensi menyebabkan wilayah tersebut terpapar lebih tinggi karena miliki interaksi yang tinggi yang disebabkan oleh proses produksi, proses perniagaan yang memiliki rata-rata 8 jam operasional.

C2 :

Ketiga wilayah ini juga merupakan pusat bisnis penopang perekonomian kota Surakarta dari transaksi hasil bumi dan laut. Fungsinya sebagai pusat bisnis menunjukkan adanya

kegiatan aktifitas perekonomian yang padat yang berpotensi menyebabkan wilayah tersebut terpapar lebih tinggi karena memiliki interaksi yang tinggi yang disebabkan oleh proses perniagaan di pasar atau kompleks pertokoan, meskipun cenderung jam operasional relatif sama. Namun pada umumnya pasar beroperasi dengan waktu yang lebih singkat dengan demikian potensi kerumunan yang disebabkan oleh kegiatan perekonomian lebih rendah dibandingkan dengan C1.

C3 :

Kedelapan wilayah ini juga merupakan pusat bisnis penopang perekonomian kota Surakarta dari transaksi komersial. Fungsinya sebagai pusat bisnis menunjukkan adanya kegiatan aktifitas perekonomian yang padat yang berpotensi menyebabkan wilayah tersebut terpapar lebih tinggi karena memiliki interaksi yang tinggi yang disebabkan oleh proses perniagaan di rumah yang dijadikan workshop, atau rumah yang dijadikan kantor atau rumah yang dijadikan toko, meskipun cenderung jam operasional relatif sama. Namun pada umumnya kegiatan perekonomian di kedelapan wilayah ini cenderung tidak berpotensi menyebabkan kerumunan orang karena profil kegiatan perekonomian yang bersifat bukan pokok kebutuhan seperti percetakan, kerajinan tangan dan sebagainya yang menunjukkan potensi terpaparnya kedelapan wilayah ini lebih rendah dari C1 dan C2.

C4 :

Merupakan wilayah yang secara umum merupakan wilayah yang memiliki kecenderungan sebagai wilayah hunian, meskipun ada kegiatan perekonomian yang tidak kalah dari wilayah di masing-masing *cluster*. Namun demikian, dengan adanya mutasi covid-19 menjadi varian delta. *Cluster* ini memiliki resiko lebih tinggi, karena warga yang telah melakukan aktifitas perekonomian seperti dagang, kerja dan bisnis kembali ke rumah untuk beristirahat yang mana hal ini dapat menyebabkan resiko terpapar lebih tinggi karena warga merasa lebih aman di rumah sehingga cenderung abai dengan kondisi pandemi.

Berdasarkan keempat data *cluster* tersebut wilayah yang memiliki zonasi resiko merah, oranye dan kuning perlu melakukan mitigasi penanganan Covid-19 salah satunya yaitu mitigasi berbasis komunal.

Mitigasi komunal merupakan metode pengendalian covid-19 pada pasien maupun orang-orang yang memiliki riwayat kontak erat dengan kasus konfirmasi, probable Covid-19.

## 5.KESIMPULAN

Kriteria zonasi covid penelitian ini menggunakan 4 level yaitu zona hijau, zona kuning, zona oranye dan zona merah menghasilkan jumlah *cluster* yaitu 4 (empat) yang C1 sebanyak 2 anggota, C2 sebanyak 3 anggota, C3 sebanyak delapan anggota dan C4 sebanyak 41 anggota. Keempat cluster tersebut menunjukkan *cluster* pertama (C1) mewakili rumah pada suatu RT di wilayah Kemlayan dan Sondakan yang dijadikan zona resiko Covid-19 sebagai zona merah karena terdapat lebih dari 5 (lima) rumah dalam satu RT, *cluster* kedua (C2) mewakili rumah pada suatu RT di wilayah Nusukan, Mojosongo dan Jebres yang dijadikan zona resiko Covid-19 sebagai zona oranye karena terdapat 3-5 rumah dalam satu RT, cluster ketiga (C3) mewakili rumah pada suatu RT di wilayah kedunglumbu, joglo, serengan, kerten, banyuanyar, jagalan, pajang dan kadipiro yang dijadikan zona resiko Covid-19 sebagai zona kuning karena terdapat 1-2 rumah dalam satu RT dan sisanya *cluster* keempat mewakili rumah dengan zonas resiko Covid-19 sebagai zona hijau (aman) namun tetap beresiko terjadi penularan Covid-19.

Perbedaan ditunjukkan bahwa ditemukan ke-tidaksesuaian antara data master dengan data

setelah *clustering*. Hal ini disebabkan karena karakteristik algoritma *k-means* yang mengevaluasi jarak berdasarkan tingkat kesamaan (*similaritas*) dengan demikian diperlukan suatu pendekatan dan atau penggunaan metode lain untuk meningkatkan akurasi hasil *clustering*.

#### DAFTAR PUSTAKA

- [1] Pemerintah Provinsi Jawa Tengah, “Instruksi Gubernur Jawa Tengah No.1 Tahun 2020.” Pemerintah Propinsi Jawa Tengah, 2020.
- [2] D. SP, “Bimbingan Teknis Aplikasi Jogo Tonggo,” 2020. .
- [3] S. Susanto and D. Suryadi, *Pengantar Data Mining Menggali Pengetahuan dari Bongkahan Data*. Yogyakarta: Penerbit ANDI, 2010.
- [4] J. Han, *Data Mining: Concepts and Techniques Second Edition*. Elsevier, 2006.
- [5] D. T. Larose, *DATA MINING METHODS AND MODELS*. John Wiley & Sons, Inc, 2006.
- [6] S. Dawaty, “K-Means Clustering,” *Univeristas Raharja*, 2020. .
- [7] T. K. P. G. T. Nasional, “Pengelompokan Kriteria Risiko COVID-19 di Daerah Berdasarkan Zonasi Warna,” *BNPB*, 2020. .
- [8] H. Hasanah, Nurmalitasari, and N. A. Sudiby, “Implementasi Data Mining Clustering untuk Mengetahui Potensi Produktifitas Kacang Tanah di Indonesia,” *NJCA (National J. Comput. Its Appl.*, vol. 5, no. 1, pp. 1–7, 2020.
- [9] N. A. Sudiby, A. Iswardani, K. Sari, and S. Suprihatiningsih, “Penerapan Data Mining Pada Jumlah Penduduk,” *Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 1, no. 3, pp. 199–207, 2020.
- [10] A. Iswardani, “Analisis Log Database Serangan Denial of Service Menggunakan Density K-Means,” Universitas Islam Indonesia, 2016.
- [11] Kusri and E. T. Luthfi, *Algoritma Data Mining*. Yogyakarta: Penerbit ANDI, 2009.
- [12] C. C. Aggarwal and C. K. Reddy, Eds., *DATA CLUSTERING Algorithms and Applications*. Chapman & Hall/CRC, 2014.
- [13] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. ANDI Offset, 2012.
- [14] D. A. I. C. Dewi and D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [15] J. Jia *et al.*, “Epidemiological Characteristics on the Clustering Nature of COVID-19 in Qingdao City, 2020: A Descriptive Analysis,” *Disaster Med. Public Health Prep.*, vol. 14, no. 5, pp. 643–647, 2020, doi: 10.1017/dmp.2020.59.

- [16] S. H. Ebrahim, Q. A. Ahmed, E. Gozzer, P. Schlagenhauf, and Z. A. Memish, "Covid-19 and community mitigation strategies in a pandemic," *BMJ*, vol. 368, no. March, pp. 1–2, 2020, doi: 10.1136/bmj.m1066.
- [17] Kementerian Sosial Republik Indonesia, *Panduan Penyiapan Fasilitas Shelter untuk karantina dan isolasi terkait COVID-19 Berbasis Komunitas*. Kementerian Sosial Republik Indonesia, 2020.