

**PENYEBARAN COVID-19 PEMPROV JAWA TIMUR  
MENGUNAKAN ANALISIS CLUSTER HIERARKI METODE  
WARD DENGAN PEARSON CORRELATION DISTANCE**

**SKRIPSI**

**OLEH  
RIZKA ALLIFATUR ROHMAH  
NIM. 16610110**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**ANALISIS CLUSTER HIERARKI METODE WARD DENGAN  
PEARSON CORRELATION DISTANCE BERDASARKAN  
PENYEBARAN COVID-19 PEMPROV JATIM**

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh  
RIZKA ALLIFATUR ROHMAH  
NIM. 16610110**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**ANALISIS CLUSTER HIERARKI METODE WARD DENGAN  
PEARSON CORRELATION DISTANCE BERDASARKAN  
PENYEBARAN COVID-19 PEMPROV JATIM**

**SKRIPSI**

**Oleh  
Rizka Allifatur Rohmah  
NIM. 16610110**

Telah Disetujui Untuk Diuji

Malang, 23 Juni 2023

Dosen Pembimbing I,



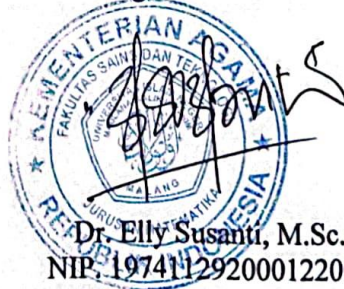
**Angga Dwi Mulyanto, M.Si.  
NIP. 198908132019031012**

Dosen Pembimbing II,



**Erna Herawati, M.Pd.  
NIDT. 19760723201802012222**

Mengetahui,  
Ketua Program Studi Matematika



**Dr. Elly Susanti, M.Sc.  
NIP. 197411292000122005**

**ANALISIS CLUSTER HIERARKI METODE WARD DENGAN  
PEARSON CORRELATION DISTANCE BERDASARKAN  
PENYEBARAN COVID-19 PEMPROV JATIM**

**SKRIPSI**

**Oleh  
Rizka Allifatur Rohmah  
NIM. 16610110**


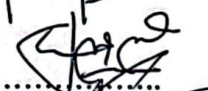
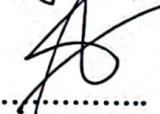

Telah Dipertahankan di Depan Dewan Penguji  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)  
Tanggal 26 Juni 2023

Ketua Penguji : Dr. Elly Susanti, M.Sc.

Anggota Penguji 1 : Ria Dhea Layla Nur Karisma, M.Si.

Anggota Penguji 2 : Angga Dwi Mulyanto, M.Si.

Anggota Penguji 3 : Erna Herawati, M.Pd

"   
.....  
  
.....  
  
.....  
  
.....

Mengetahui,  
Ketua Program Studi Matematika

  
Dr. Elly Susanti, M.Sc.  
NIP. 19741129 200012 2 005

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Rizka Allifatur Rohmah

NIM : 16610110

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Persebaran Covid-19 Pemprov Jawa Timur menggunakan analisis *cluster* Hierarki Metode *Ward* dengan Pearson Correlation distance.

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 26 Juni 2023

Yang membuat pernyataan,



Rizka Allifatur Rohmah  
NIM. 16610110

## **MOTTO**

**Dan (ingatlah) ketika Musa berkata kepada muridnya: “Aku tidak akan berhenti (berjalan) sebelum sampai ke pertemuan dua buah lautan , atau aku akan berjalan sampai bertahun-tahun” (QS. Al-Kahfi [18]:60)**

**”Tak peduli apa yang terjadi, Teguhlah seperti Musa, teruslah berjalan hingga mencapai apa yang engkau cita-citakan”**

## KATA PENGANTAR

*Assalamualaikum Warahmatullahi Wabarakatuh.*

Segala puji bagi Allah SWT yang selalu melimpahkan rahmat, taufik dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Cluster Hierarki Metode Ward dengan Pearson Correlation Distance berdasarkan Penyebaran Covid-19 Pemprov Jatim” sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang Matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Sholawat dan salamsemoga tetap tercurahkan kepada junjungan kita Nabi Muhammad SAW yang telah menuntun manusia dari jalan kegelapan menuju ke jalan yang terang benderang yaitu Islam.

Dalam penyusunan skripsi ini tidak lepas dari petunjuk dan bimbingan serta masukan dari berbagai pihak. Untuk itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
3. Dr. Elly Susanti, M.Sc. selaku ketua Program Studi Matematika, Universitas Islam Negeri Maulana Malik Ibrahim.
4. Angga Dwi Mulyanto, M.Si, selaku dosen pembimbing I yang begitu sabar dan teliti dalam membimbing serta memberikan arahan kepada penulis untuk menyelesaikan skripsi ini.

5. Erna Herawati, M.Pd, selaku dosen pembimbing II yang begitu sabar dan teliti dalam membimbing serta memberikan arahan kepada penulis untuk menyelesaikan skripsi ini.
6. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim yang telah memberikan segala ilmu dan bimbingannya.
7. Kedua orang tua Ayah dan Ibu serta Adik Penulis dan Suami, yang telah memberikan motivasi, semangat serta dukungan dan segala pengorbanan baik moril maupun materil.
8. Dan kepada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini.

*Wassalamualaikum warahmatullahi Wabarakatuh*

Malang, 26 Juni 2023

Penulis



## DAFTAR ISI

HALAMAN PENGANTAR .....	i
HALAMAN PERSETUJUAN .....	ii
HALAMAN PENGESAHAN .....	iii
PERNYATAAN KEASLIAN TULISAN .....	iv
MOTTO .....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL .....	x
DAFTAR GAMBAR.....	xi
ABSTRAK .....	xii
ABSTRACT .....	xiii
مستخلص البحث.....	xiv
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	5
1.6 Definisi Istilah .....	5
<b>BAB II KAJIAN PUSTAKA .....</b>	<b>7</b>
2.1 Analisis Cluster .....	7
2.1.1 Proses Analisis <i>Cluster</i> .....	9
2.1.2 Metode Pengelompokan .....	10
2.2 Principal Component Analysis.....	17
2.3 Validasi Cluster .....	17
2.3.1 Indeks Dunn .....	18
2.3.2 Indeks Silhouette.....	19
2.3.3 Indeks Connectivity .....	19
2.4 Uji Asumsi Analisis Cluster.....	19
2.4.1 Uji Korelasi .....	19
2.5 Covid-19.....	20
2.6 Kajian Integrasi topik dalam Islam.....	22
2.6.1 Kajian Integrasi Islam tentang wabah .....	22
2.6.2 Kajian Integrasi Islam tentang <i>Cluster</i> .....	24
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>26</b>
3.1 Pendekatan Penelitian .....	26
3.2 Sumber Data.....	26
3.3 Variabel Penelitian .....	26
3.4 Metode Analisis Data .....	26
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>28</b>
4.1 Analisis Deskriptif.....	28
4.2 Uji Asumsi Cluster.....	33
4.2.1 Uji Multikolinieritas.....	33
4.2.2 Principal Component Analysis (PCA) .....	34
4.2.3 Uji Kaiser-Meyer Olkin .....	35
4.3 Penentuan jumlah <i>Cluster</i> Optimum.....	35

4.4	Clustering Hierarchical .....	36
4.5	Interpretasi Metode Terbaik .....	39
4.6	Integrasi Al-qur'an dan Cluster .....	39
<b>BAB V PENUTUP .....</b>		<b>45</b>
5.1	Kesimpulan.....	45
5.2	Saran.....	46
<b>DAFTAR PUSTAKA .....</b>		<b>47</b>
<b>LAMPIRAN.....</b>		<b>48</b>
<b>RIWAYAT HIDUP .....</b>		<b>58</b>

## DAFTAR TABEL

Tabel 4.1	Matriks Korelasi .....	33
Tabel 4.2	<i>Eigen Values</i> .....	34
Tabel 4.3	Hasil PCA.....	35
Tabel 4.4	Koefisien Korelasi.....	35
Tabel 4.5	Penentuan Jumlah Cluster Optimum.....	36
Tabel 4.6	<i>Agglomerative coefficient</i> .....	37
Tabel 4.7	Anggota <i>Cluster</i> .....	38
Tabel 4.8	Profilisasi <i>Cluster</i> Terbaik .....	39

## DAFTAR GAMBAR

Gambar 3.1	<i>Flowchart</i> Penelitian .....	27
Gambar 4.1	Kasus Probable Covid-19 Kabupaten/Kota di Jawa Timur.....	28
Gambar 4.2	Kasus Konfirmasi Covid-19 Kabupaten/Kota di Jawa Timur..	29
Gambar 4.3	Kasus Meninggal Covid-19 Kabupaten/Kota di Jawa Timur...	31
Gambar 4.4	Kasus Aktif Covid-19 Kabupaten/Kota di Jawa Timur .....	32
Gambar 4.5	Dendogram Ward .....	37

## ABSTRAK

Rizka Allifatur Rohmah. 2023. **Penyebaran Covid-19 Pemprov Jawa Timur Menggunakan Analisis Cluster Hierarki Metode Ward Dengan Pearson Correlation Distance**. Tugas akhir/skripsi. Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Angga Dwi Mulyanto, M.Si (II) Erna Herawati M.Pd.

**Kata Kunci:** *Cluster Hierarki, Covid-19, Ward*

*Coronavirus disease* atau COVID-19 adalah penyakit menular yang disebabkan oleh infeksi *Severe Acute Respiratory Syndrome Coronavirus 2* (SARS-CoV-2). Hal sederhana yang paling efektif untuk memutus rantai penyebaran virus adalah meningkatkan daya tahan tubuh dengan asupan makanan yang sehat, rajin mencuci tangan, dan menggunakan masker saat berada di luar rumah. Metode yang digunakan dalam penelitian ini adalah Metode *ward*. Metode *ward* adalah proses pengelompokan menggunakan pendekatan analisis varian untuk menghitung jarak antar cluster dengan cara meminimumkan jumlah kuadrat. *Clustering Hierarchical* adalah suatu metode analisis cluster yang dilakukan secara bertahap dan bertingkat sehingga membentuk tingkatan seperti pohon. Hasil dari metode ini dapat disajikan dengan dendrogram. Penentuan metode terbaik pada *cluster hierarchical* dapat ditentukan dengan *agglomerative coefficient*. didapatkan bahwa metode *hierarchical* yang menghasilkan *cluster* terbaik yaitu metode *ward* karena nilai *agglomerative coefficient* lebih besar dibandingkan dengan metode *hierarchical* lainnya. Berdasarkan hasil *cluster* menggunakan algoritma *hierarchical* pada data mengenai kasus covid-19 berdasarkan Kabupaten/Kota di Provinsi Jawa Timur didapatkan terdapat 2 cluster yaitu diantaranya, *cluster* 1 terdiri dari 1 Kabupaten/Kota, dan *cluster* 2 terdiri dari 37 Kabupaten/Kota di Provinsi Jawa Timur.

## ABSTRACT

Rizka Allifatur Rohmah. 2023. **The Spread of COVID-19 in East Java Provincial Government Using Hierarchical Cluster Analysis Ward Method with Pearson Correlation Distance.** Thesis. Departement of Mathematics, Faculty of Science and Technology, Maulana Malik Ibrahim Malang. Advisor: (I) Angga Dwi Mulyanto, M.Si (II) Erna Herawati M.Pd.

**Keyword: Clustering Hierarchical, Covid-19, Ward**

Coronavirus disease, or COVID-19, is an infectious disease caused by infection with Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The most effective simple thing to break the chain of spreading the virus is to increase the body's resistance by eating a healthy diet, washing hands diligently, and using a mask when outside the home. The method used in this research is the ward method. The ward method is a grouping process using the analysis of variance approach to calculate the distance between clusters by minimizing the sum of squares. Hierarchical clustering is a cluster analysis method that is carried out in stages so as to form levels like a tree. The results of this method can be presented with a dendrogram. The best method in the cluster hierarchy can be determined by agglomerative coefficients. It was found that the hierarchical method that produces the best clusters is the ward method because the agglomerative coefficient value is greater than the other hierarchical methods. Based on the results of clusters using a hierarchical algorithm on data regarding covid-19 cases based on districts/cities in East Java Province. It was found that there were two clusters, namely, cluster 1 consisting of one Regency/City, and cluster 2 consisting of 37 Regencies/Cities in East Java Province.

## مستخلص البحث

رزكا اليفة رحمة. 2023. انتشار حكومة مقاطعة جاوة الشرقية Covid-19 باستخدام أسلوب وارد التحليل العنقودي الهرمي مع مسافة ارتباط بيرسون. البحث العلمي. قسم الرياضيات. كلية العلوم والتكنولوجيا, جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (1) أنجا دوي موليانطا ماجستير. (2) إيرنا هيراواتيم اجستير.

### الكلمات الدالة: *Clustering Hierarchical, Covid-19, Ward*

مرض فيروس كورونا أو COVID-19 هو مرض معد تسببه العدوى بفيروس كورونا ٢ المسبب للمتلازمة التنفسية الحادة الوخيمة (SARS-CoV-2). أبسط شيء أكثر فعالية لكسر سلسلة انتشار الفيروس هو زيادة القدرة على التحمل عن طريق تناول الأطعمة الصحية ، وغسل اليدين بجد ، واستخدام الأقنعة عندما تكون خارج المنزل. الطريقة المستخدمة في هذه الدراسة هي Ward. Ward هي عملية تجميع باستخدام نهج تحليل التباين لحساب المسافات بين المجموعات عن طريق تقليل عدد المربعات. Clustering Hierarchical هو طريقة للتحليل العنقودي يتم تنفيذها تدريجياً وطبقياً لتشكيل مستويات تشبه الشجرة. يمكن تقديم نتيجة هذه الطريقة مع مخطط شجري. يمكن تحديد أفضل طريقة على Clustering Hierarchical بواسطة المعاملات التكتلية. وجد أن الطريقة الهرمية التي تنتج أفضل مجموعة هي طريقة Ward لأن قيمة المعامل التكتلي أكبر من الطرق الهرمية الأخرى. استناداً إلى نتائج المجموعة باستخدام خوارزمية هرمية على البيانات المتعلقة بحالات covid-19 بناء على المنطقة أو المدينة في مقاطعة جاوة الشرقية ، وجد أن هناك مجموعتين ، بما في ذلك المجموعة 1 التي تتكون من 1 المنطقة أو المدينة ، والمجموعة 2 التي تتكون من 37 المنطقة أو المدينة في مقاطعة جاوة الشرقية.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Analisis *cluster* merupakan salah satu cabang ilmu dalam statistik, yang digunakan untuk mengelompokkan objek-objek tertentu berdasarkan kesamaan karakteristik data pada kelompok. Analisis *cluster* secara umum memiliki dua metode yaitu, metode hierarki dan non hierarki. Pengelompokan metode hierarki digunakan ketika tidak terdapat informasi tentang jumlah *cluster*, sedangkan dalam pengelompokan metode non hierarki digunakan ketika jumlah *cluster* diketahui atau ditentukan sebagai bagian prosedur pengelompokan (John & Winchern, 2007).

Untuk mengidentifikasi kelompok kabupaten atau kota di Provinsi Jawa Timur , perlu dilakukan pengelompokan jumlah penyebaran COVID-19 dalam suatu daerah pengelompokan dapat dilakukan dengan analisis *clustering*. Analisis *cluster* adalah salah satu teknik multivariat yang bertujuan mengklasifikasi suatu objek-objek ke dalam suatu kelompok-kelompok yang berbeda antara kelompok satu dengan yang lainnya. Objek-objek yang telah diklasifikasikan dalam satu *cluster* merupakan objek-objek yang memiliki kedekatan jarak relatif sama dengan objek lainnya (Narimawati, 2008).

Metode yang digunakan dalam penelitian ini adalah Metode *Ward*. Metode *Ward* merupakan salah satu metode dalam analisis *cluster* hierarki. Metode *Ward* adalah proses pengelompokan menggunakan pendekatan analisis varian untuk menghitung jarak antar *cluster* dengan cara meminimumkan jumlah kuadrat.



*Coronavirus* disease atau sering kira sebut COVID-19 adalah penyakit menular yang disebabkan oleh infeksi Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Pada 31 Desember 2019 kasus ini berkembang pesat di Kota Wuhan Cina. Pada tanggal 30 Januari 2020 WHO menetapkan Kedaruratan Kesehatan Masyarakat yang meresahkan Dunia (KKMMD) dan pada tanggal 11 Maret 2020 WHO menetapkan COVID-19 sebagai pandemi. Hal sederhana yang paling efektif untuk memutus rantai penyebaran virus adalah tetap di rumah saja. Pencegahan lain adalah dengan cara meningkatkan daya tahan tubuh dengan asupan makanan yang sehat, rajin mencuci tangan, melakukan olahraga, menggunakan masker saat berada di luar rumah, dan istirahat yang cukup.

Dalam Al-qur'an surat Ar-ra'd ayat 11 disebutkan bahwa:

لَهُ مُعَقَّبَاتٌ مِّنْ بَيْنِ يَدَيْهِ وَمِنْ خَلْفِهِ يَحْفَظُونَهُ مِنْ أَمْرِ اللَّهِ إِنَّ اللَّهَ لَا يُغَيِّرُ مَا بِقَوْمٍ حَتَّىٰ يُغَيِّرُوا مَا بِأَنْفُسِهِمْ وَإِذَا أَرَادَ اللَّهُ بِقَوْمٍ سُوءًا فَلَا مَرَدَّ لَهُ وَمَا لَهُمْ مِّنْ دُونِهِ مِنْ وَالٍ (١١)

Artinya:

*“Baginya (manusia) ada (malaikat-malaikat) yang menyertainya secara bergiliran dari depan dan belakangnya yang menjaganya atas perintah Allah. Sesungguhnya Allah tidak mengubah keadaan suatu kaum hingga mereka mengubah apa yang ada pada diri mereka. Apabila Allah menghendaki keburukan terhadap suatu kaum, tidak ada yang dapat menolaknya, dan sekali-kali tidak ada pelindung bagi mereka selain Dia.”*

Ayat di atas menjelaskan bahwa Allah tidak akan mengubah keadaan suatu kaum sebelum mereka mengubah keadaan mereka sendiri, salah satu bentuk ikhtiar kita menghadapi Covid-19 yaitu dengan cara mematuhi Protokol Kesehatan, tetap berada di rumah saja, menjahui kerumunan, dan melakukan vaksinasi. Ikhtiar lain dalam menghadapi Covid-19 yaitu dengan penerapan Lockdown seperti hadits Rasulullah “Jika kamu mendengar wabah disuatu

wilayah, maka janganlah kalian memasukinya. Tapi jika terjadi wabah di tempat kamu berada maka jangan tinggalkan”. Dari penjelasan hadits tersebut disimpulkan bahwa salah satu ikhtiar kita yaitu dengan Lockdown atau Karantina Wilayah.

Sebagaimana telah disebutkan didalam surat At-taghabun ayat 11 berbunyi:

مَا أَصَابَ مِنْ مُصِيبَةٍ إِلَّا بِإِذْنِ اللَّهِ وَمَنْ يُؤْمِنْ بِاللَّهِ يَهْدِ اللَّهُ قَلْبَهُ وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ

Artinya :

*“Tidak ada suatu musibah menimpa (seseorang, kecuali dengan izin Allah, dan barangsiapa beriman kepada Allah, niscaya Allah akan memberi petunjuk kepada hatinya. Dan Allah mengetahui segala sesuatu”.*

Berdasarkan ayat diatas dijelaskan bahwa segala musibah datangnya adalah dari Allah, segala sesuatu yang menimpa manusia adalah qada’ dan qodar Allah yang sudah dikehendakinya, dalam hal ini hendaknya manusia tidak menyesali dan merasa kecewa apabila menemui hal-hal yang tidak sesuai dengan keinginannya, karena semua hal itu adalah diluar kemampuan kita, sesungguhnya Allah akan memberikan petunjuk kepada prang yang beriman dan melapangkan dadanya dengan menerima segala apa yang terjadi pada dirinya, karena yakin bahwa semua itu dari Allah. Usaha terakhir dalam menghadapi Covid-19 ini setelah sabar dan ikhtiar adalah tawakal. Dalam Islam kita tidak hanya diwajibkan untuk ikhtiar dan tawakal saja melainkan Islam memerintahkan kita agar melaksanakan keduanya secara bersama. Brangsiapa beriman kepada Allah saat tertimpa musibah, kemudian dia menyadari bahwa semua adalah takdir Allah, kemudian dia bersabar dan bertawakal maka Allah akan memberikan petunjuk dan menggantikan sesuatu yang lebih baik (Ad-Dimasyqi, 2000).

Pada penelitian sebelumnya membahas tentang analisis cluster yang telah dilakukan Sukmawati (2017) yang berjudul “Analisis *cluster* dengan metode hierarki untuk pengelompokan kabupaten/kota di provinsi Sulawesi selatan berdasarkan indikator makro ekonomi” dengan menggunakan metode *ward*. Hasil yang didapatkan adalah empat kelompok kabupaten/kota memiliki kemiripan ekonomi. Berdasarkan uraian di atas, hal tersebut yang melatar belakangi penulis untuk melakukan penelitian dengan judul “Penyebaran Covid-19 Pemprov Jawa Timur Menggunakan Analisis *Cluster* Hierarki Metode *Ward* dengan Pearson Correlation Distance”.

## **1.2 Rumusan Masalah**

Berdasarkan pada uraian latar belakang, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana hasil pengelompokan data menggunakan metode *ward* pada sebaran COVID-19 di Pemprov Jawa Timur?
2. Bagaimana karakteristik dari setiap kelompok yang terbentuk?

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah:

1. Mengetahui hasil pengelompokan menggunakan jarak *pearson correlation* pada analisis cluster hirarki pada sebaran COVID-19 di Jawa Timur.
2. Mengetahui karakteristik pada setiap kelompok cluster yang terbentuk.

## **1.4 Batasan Masalah**

Untuk membatasi masalah agar sesuai dengan apa yang dimaksudkan dan tidak menimbulkan masalah yang baru, maka peneliti memberikan batasan masalah:

1. Metode yang digunakan dalam penelitian ini yaitu *Metode Ward*
2. Metode jarak yang digunakan adalah *Pearson Correlation distance*
3. Data yang digunakan dalam penelitian ini adalah data sekunder mengenai COVID-19 yang diperoleh dari data persebaran COVID-19 Jawa Timur

### **1.5 Manfaat Penelitian**

Adapun manfaat dalam penelitian ini adalah:

1. Menambah wawasan tentang analisis cluster hierarki jika digunakan dalam metode *ward* pada penyebaran covid-19
2. Menambah wawasan tentang aplikasi analisis cluster dalam mengidentifikasi penyebaran covid-19 dengan metode *ward*.
3. Memudahkan pada proses pengelompokkan data berdasarkan karakteristik yang sama.

### **1.6 Definisi Istilah**

1. *Pearson Correlation Distance*

Teknik untuk mengukur kekuatan hubungan antara variabel satu dengan lainnya dan juga dapat mengetahui bentuk hubungan antar variabel.

2. *Analisis cluster*

Salah satu teknik analisis yang bertujuan untuk mengelompokkan objek-objek kedalam sesuatu yang mempunyai kesamaan karakteristik antar objek satu dengan yang lainnya.

3. *Pearson Correlation*

Teknik untuk mengukur kekuatan hubungan antara variabel satu dengan lainnya dan juga dapat mengetahui bentuk hubungan antar variabel.

#### 4. Metode *Ward*

Metode varians yang bertujuan untuk memperoleh *cluster* yang memiliki varians internal *cluster* yang sekecil mungkin.

## BAB II

### KAJIAN PUSTAKA

#### 2.1 Analisis Cluster

Analisis *cluster* merupakan salah satu teknik multivariat yang memiliki tujuan untuk mengelompokkan suatu objek menjadi suatu kelompok yang berbeda dengan kelompok lainnya. Objek yang telah dikategorikan menjadi satu *cluster* adalah objek yang mempunyai kedekatan jarak relatif dengan objek yang lainnya (Narimawati, 2008).

*Cluster* juga dapat diartikan kelompok, pada dasarnya analisis *cluster* juga akan menghasilkan kelompok-kelompok. Analisis *cluster* didasari dengan objek yang memiliki kemiripan dengan objek lain. Oleh karena itu dalam pengelompokkan mempunyai karakteristik yang serupa dalam satu *cluster* atau lebih (Santoso, 2018).

Analisis *cluster* merupakan suatu teknik multivariate, akan tetapi dalam analisis ini berbeda dengan analisis multivariate yang lain. Pada analisis *cluster* variat diartikan sebagai sejumlah variabel yang digunakan sebagai objek untuk membandingkan dengan variabel lainnya, sedangkan pada variat lainnya diartikan sebagai kombinasi linier berbagai variabel. Maka dalam analisis *cluster*, tidak dilakukan pencarian nilai variat secara empiris, sebagaimana yang dilakukan multivariat yang lainnya akan tetapi analisis *cluster* mempunyai tujuan yaitu untuk menempatkan sekelompok objek ke dalam dua atau lebih *cluster* berdasarkan kesamaan objek atas dasar berbagai karakteristik (Nafisah & Candra, 2017).

Analisis *cluster* disebut juga dengan kelas teknik yang digunakan untuk mengklasifikasi objek atau kasus dalam kelompok yang relatif homogen yang

disebut *cluster*. Objek dalam kelompok cenderung mempunyai kemiripan atau sangat berbeda dengan *cluster* lainnya. Analisis *cluster* juga disebut dengan klasifikasi atau taksonomi numerik (Supranto, 2004).

Analisis *cluster* adalah pengelompokan data ke dalam sejumlah kelompok berdasarkan kemiripan karakteristik data pada kelompok, Data atau objek yang masuk dalam batas kemiripan dengan kelompoknya akan bergabung menjadi satu kelompok, dan kemudian akan terpisah dalam kelompok yang berbeda jika keluar dari batas kemiripan dengan kelompok tersebut (Santosa, 2007).

Tujuan dari pengklasteran tersebut adalah untuk mengelompokkan objek berdasarkan kesamaan karakteristik dari objek-objek tersebut. Langkah yang digunakan dalam mengukur kesamaan yaitu dengan jarak antar objek. Semakin besar nilai kesamaan objek maka semakin besar perbedaan antar objek tersebut. Analisis *cluster* juga diterapkan dalam berbagai ilmu misalnya dalam dunia kedokteran,biologi,manajemen psikologi dan sebagainya (Santoso, 2018).

Berikut adalah ciri-ciri dari *cluster cluster* yang baik:

1. Homogenitas (kesamaan) yang tinggi antar anggota dalam satu kelompok.
2. Heterogenitas (perbedaan) yang sangat tinggi antara satu kelompok dengan kelompok lainnya.

Selain homogenitas dan heterogenitas,terdapat asumsi tentang non-multikolinearitas. Multikolinearitas atau kolinearitas ganda adalah hubungan linear atau korelasi antar variabel-variabel bebas, gejala multilolinearitas dapat diketahui dengan beberapa cara salah satunya yaitu dengan korelasi person (Lina & Eni, 2011).

### 2.3.1 Proses Analisis *Cluster*

Dalam analisis *cluster* terdiri dari beberapa tahap dasar, yaitu

(Supranto, 2004):

1. Merumuskan Masalah

Merumuskan masalah dalam analisis *cluster* adalah hal yang sangat penting yaitu dalam proses pemilihan variabel yang akan digunakan sebagai objek dalam pengklasteran.

2. Memilih Similaritas atau Ukuran Jarak.

Tujuan dari penklasteran yaitu untuk mengelompokkan objek yang mempunyai kesamaan, maka ukuran sangat diperlukan untuk mengetahui kemiripan atau perbedaan setiap objek. Pendekatan paling sederhana yaitu dengan mengukur kemiripan dengan jarak antara masing-masing objek. Jika jarak semakin dekat maka cenderung memiliki kemiripan daripada jarak jauh.

3. Standarisasi Data

Proses standarisasi dilakukan apabila terdapat variabel yang memiliki satuan yang berbeda. Karena dalam hasil perhitungan *cluster* mengakibatkan hasil yang tidak valid. Proses standarisasi dilakukan dengan transformasi yaitu standarisasi data asli sebelum dianalisis lebih lanjut proses tersebut dilakukan dengan menggunakan Z-skor.

4. Menentukan Proses Pengklusteran

Setelah melakukan standarisasi data dan menentukan metode *cluster* (misalnya dipilih metode *Ward*) proses selanjutnya yaitu pengelompokan data.



### 5. Interpretasi Hasil *Cluster* Yang Telah Terbentuk

Setelah proses *cluster* terbentuk maka proses selanjutnya adalah melakukan interpretasi pada *cluster* yang sudah terbentuk. Langkah interpretasi yaitu dengan pengujian tiap *cluster* untuk menamai dan menandai kealamian *cluster*. Proses ini dimulai dengan suatu ukuran yaitu dengan *Centroid Cluster*.

### 6. Melakukan Proses Validasi dan Profiling *Cluster*

Proses validasi ini bertujuan untuk menjamin bahwa hasil dari *cluster* dapat mewakili populasi dan dapat digeneralisasikan pada objek lain. Pendekatan ini dilakukan dengan membandingkan solusi *cluster* dan melalui hasil *korespondensi*. Jika *cluster* yang terbentuk sudah valid maka proses selanjutnya profiling dengan tujuan menjelaskan hasil *cluster* berdasarkan profil tertentu. Hasil *cluster* dapat digunakan dalam berbagai kepentingan sesuai dengan bahan yang akan diteliti.

## 2.3.2 Metode Pengelompokan

Metode *clustering* secara umum dapat dibagi menjadi dua metode yaitu *cluster* hierarki dan non-hierarki. Metode hierarki adalah metode pengelompokan yang bertahap dan berstruktur berdasarkan kemiripan sifat antar objek. Sedangkan metode non-hierarki adalah dimulai dengan memilih nilai *cluster* awal sesuai dengan nilai yang diinginkan.

Metode Hierarki atau *Hierarchical Method* adalah pengelompokan dengan dua objek atau lebih yang mempunyai objek paling dekat. Kemudian dilakukan pengelompokan ke objek lain yang kedekatan kedua. Proses ini dilakukan secara bertahap dan bertingkat dan membentuk seperti struktur sebuah pohon. Hasil dari pengklasteran tersebut disajikan dalam bentuk dendogram. Dendogram merupakan representasi visual dari langkah-langkah

dalam *analisis cluster* yang menunjukkan bagaimana terbentuknya *cluster* dan nilai koefisien jarak pada setiap langkah (Simamora, 2005).

Ada beberapa cara untuk mengukur jarak (*distance*) dalam analisis *cluster*, antara lain (Mongi, 2015):

1. *Euclidean Distance*

Jarak ini sering digunakan dalam analisis cluster jarak ini digunakan apabila variabel pada jarak ini berskala kontinu. Pada jarak ini pengukurannya dilakukan dengan menghitung akar kuadrat dari penjumlahan kuadrat selisih dari nilai masing-masing peubah. Jarak euclid dapat dirumuskan sebagai berikut:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (2.1)$$

2. *Manhattan Distance*

*Manhattan distance* atau *City blok distance*, jarak ini merupakan bentuk umum dari *euclied distance*. Jarak ini digunakan untuk menghitung perbedaan absolut (mutlak) antara koordinat sepasang objek. Jarak Manhattan dapat dirumuskan sebagai berikut:

$$d(i, k) = \sum_{j=1}^n |x_{ij} - x_{kj}| \quad (2.2)$$

Dengan:

$d_{i,k}$  : Jarak antara data ke- $i$  dengan data ke- $k$

$x_{i,j}$  : Nilai data ke- $i$  pada peubah ke- $j$

$x_{k,j}$  : Nilai data ke- $k$  pada peubah ke- $j$

$n$  : Jumlah peubah yang diamati

3. *Mahalonobis Distance*

4. Jarak *Mahalonobis* ini digunakan untuk menghilangkan atau mengatasi variabel yang memiliki skala berbeda. Jarak *Mahalonobis* dapat dirumuskan sebagai berikut:

$$d_{ik} = \sqrt{(x_{ij} - x_{kj})'S^{-1} + (x_{ij} - x_{kj})} \quad (2.3)$$

dengan:

$d_{ik}$  : Jarak antara data ke- $i$  dengan data ke- $k$

$x_{i,j}$  : Nilai data ke- $i$  pada peubah ke- $j$

$x_{k,j}$  : Nilai data ke- $k$  pada peubah ke- $j$

$S$  : Matriks covarian

5. *Log-Likelihood distance*

Jarak ini digunakan pada peubah berskala kontinu dan kategorik. Jarak ini dapat dirumuskan sebagai berikut:

$$d(j, s) = \xi_j + \xi_s - \xi_{(js)} \quad (2.4)$$

dengan:

$$\xi_j = -N \left( \sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{jk}^2) - \sum_{k=1}^{K^B} \sum_{i=1}^{L^k} \frac{N_{jkl}}{N_j} \log \left( \frac{N_{jkl}}{N_j} \right) \right)$$

$$\xi_j = -N \left( \sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{jk}^2) - \sum_{k=1}^{K^B} \sum_{i=1}^{L^k} \frac{N_{skl}}{N_j} \log \left( \frac{N_{skl}}{N_j} \right) \right)$$

$$\xi_{(js)} = -N \left( \sum_{k=1}^{K^A} \frac{1}{2} \log(\hat{\sigma}_k^2 + \hat{\sigma}_{(js)k}^2) - \sum_{k=1}^{K^B} \sum_{l=1}^{L^k} \frac{N_{(js)kl}}{N_j} \log \left( \frac{N_{(js)kl}}{N_j} \right) \right)$$

Dengan:

- $N$  : Jumlah objek
- $N_j$  : Jumlah objek gerombol didalam
- $N_{jkl}$  : Jumlah objek digerombol  $j$  untuk peubah kategori ke-  $k$  dengan kategorik ke-  $l$
- $\hat{\sigma}_k^2$  : Ragam dugaan untuk peubah kontinu ke- $k$  untuk keseluruhan objek
- $\hat{\sigma}_{jk}^2$  : Ragam dugaan untuk peubah kontinu ke- $k$  untuk keseluruhan objek dalam gerombol  $j$
- $K^A$  : Banyaknya peubah kontinu
- $K^B$  : Banyaknya peubah kategorik
- $L^k$  : Banyaknya kategori untuk peubah kategorik ke- $k$

#### 6. *Chi-square Distance*

Jarak *chi-square* ini menghitung antara profil dalam jarak euclidean terboboti menggunakan invers proporsi rata-rata sebagai bobot. Misal  $c_j$  menunjukkan elemen ke- $j$  dari rata-rata profil, yang merupakan kelimpahan proporsi peubah  $j$  dalam seluruh kumpulan data. Maka jarak

*chi-square* dinotasikan  $\chi$  antara dua objek dan profil  $x = [x_1, x_2, \dots, x_j]$

dan  $y = [y_1, y_2, \dots, y_j]$  didefinisikan sebagai:

$$\chi_{x,y} = \sqrt{\sum_{j=1}^J \frac{1}{c_j} (x_j - y_j)^2} \quad (2.5)$$

#### 7. *Pearson Correlation Distance*

Jarak ini merupakan jarak yang mengevaluasi sejauh mana objek memiliki hubungan satu sama lain. Jarak ini digunakan apabila terdapat multikolinearitas pada peubah. Jarak ini dapat dirumuskan sebagai berikut:

$$d(i,j) = 1 - r_{ij} \quad (2.6)$$

dengan

$d(i,j)$  : jarak korelasi pearson

$r_{ij}$  : koefisien korelasi

Koefisien korelasi dapat dihitung dengan rumus:

$$r_{ij} = \frac{cov(X_i, X_j)}{\sqrt{var(X_i)var(X_j)}} \quad (2.7)$$

dengan:

$cov(X_i, X_j)$  = peragam variabel  $(X_i, X_j)$

$var(X_i)$  = ragam variabel ke- $i$

$var(X_j)$  = ragam variabel ke- $j$

Pada analisis cluster hierarki terdapat dua metode pengelompokan yaitu

1. Metode *Agglomerative*

Metode ini dimulai dengan mengelompokkan objek yang mempunyai jarak terdekat. Proses selanjutnya menggabungkan antar kelompok yang memiliki jarak paling dekat, selanjutnya dilakukan langkah yang sama. Metode *Agglomerative* mempunyai beberapa macam yaitu:

- a. Metode *Single Linkage*, Metode ini akan mengelompokkan dua objek yang mempunyai jarak terdekat lebih dahulu atau aturan tetangga dekat (*nearest neighbour rule I*). Sehingga dapat dirumuskan sebagai berikut:

$$d_{(x,y)z} = \min (d_{xz} , d_{yz}) \quad (2.8)$$

- b. Metode *Complete Linkage*, Metode ini akan mengelompokkan dua objek yang mempunyai jarak terjauh dahulu. Selanjutnya dilanjutkan untuk jarak variabel yang makin dekat. Sehingga dapat dirumuskan sebagai berikut:

$$d_{(x,y)z} = \max (d_{xz} , d_{yz}) \quad (2.9)$$

- c. Metode *Average Linkage*, Metode ini akan mengelompokkan dua objek berdasarkan jarak rata-rata yang mana jarak tersebut dihitung pada cluster masing-masing. Jika dalam kelompok x dan y mempunyai jarak  $d_{xy}$  maka harus dicari rata-rata jarak xz dan yz maka dapat dirumuskan sebagai berikut:

$$d_{(xy)z} = \frac{n_x}{n_x+n_y} d_{xz} + \frac{n_y}{n_x+n_y} d_{yz} \quad (2.10)$$

dengan:

$n_x$  = jumlah individu pada kelompok x

$n_y$  = jumlah individu pada kelompok y

## d. Median

Jarak antar kelompok dirumuskan dengan:

$$d_{(xy)z} = \frac{1}{2}(d_{xz} + d_{yz}) - \frac{1}{4}d_{xy} \quad (2.11)$$

e. Metode *Ward's*

Metode *Ward's* ini didasarkan pada nilai terkecil varian dalam kelompok. Maka dapat menggunakan rumus sebagai berikut:

$$d_{(xy)z} = \frac{(n_x + n_y)d_{xz} + (n_y + n_z)d_{yz} - n_z d_{xy}}{n_x + n_y + n_z} \quad (2.12)$$

Dengan

$n_x$  = jumlah individu pada kelompok  $x$

$n_y$  = jumlah individu pada kelompok  $y$

$n_z$  = jumlah individu pada kelompok  $z$

f. Metode *Centroid*

Metode *Centroid* ini adalah rata-rata semua objek didalam cluster. Metode ini jarak dalam cluster adalah jarak antar centroid. Centroid ini dihitung ketika setiap kali objek digabungkan, sehingga setiap kali anggotanya bertambah maka centroidnya akan berubah. Jarak antar cluster pada metode ini adalah jarak antar centroid (Johnson & Wichern, 1992).

2. Metode *Devisive*

Metode *Devisive* merupakan kebalikan dari metode *agglomerative*.

Metode *Devisive* ini dimulai dengan mengumpulkan semua objek menjadi satu kluster. Kemudian objek dipisah secara bertahap dengan

cluster yang berbeda sampai semua objek menjadi kluster masing-masing (Simamora,2015).

## **2.2 Principal Component Analysis**

Pengurangan dimensi (*dimension reduction*) yang terdapat pada sekelompok data digunakan untuk mengurangi beberapa variable yang tidak berelasi antarsatu variabel dengan variabel lainnya dalam satu kelompok klaster. Tujuan dari pengurangan dimensi adalah untuk mendapatkan variabel – variabel yang optimal yang dapat membentuk klaster yang diinginkan. Salah satu metode pengurangan dimensi adalah *Principal Component Analysis*.

Principal Component Analysis menurut Lindsay (2002) adalah sebuah metode untuk mengidentifikasi pola – pola yang terdapat dalam sebuah data dan menyatakannya dalam sebuah cara untuk menentukan kemiripan dan perbedaan yang dimiliki oleh data tersebut. Salah satu keunggulan yang dapat ditemukan dalam *Principal Component Analysis* adalah dengan melakukan metode ini dapat mengurangi jumlah dimensi yang terdapat dalam satu pola tanpa mengurangi informasi yang terdapat dalam data tersebut. Oleh karena itu, *Principal Component Analysis* sangat diperlukan untuk membantu reduksi terhadap pola – pola yang ada dalam suatu klaster.

## **2.3 Validasi Cluster**

Validasi cluster adalah prosedur yang mengevaluasi hasil analisis cluster secara kuantitatif dan objektif. Menurut Jain dan Dubes (1988), terdapat tiga pendekatan untuk mengeksplorasi validasi cluster, antara lain:



1. Validasi eksternal, mengevaluasi hasil dari metode clustering berdasarkan praspesifikasi struktur yang diterima dari sebuah data yang mencerminkan intuisi pengguna tentang struktur clustering dari data.
2. Validasi internal, mengevaluasi hasil clustering dalam konsep kuantitatif yang didapat dari data.
3. Validasi relatif, membandingkan sebuah struktur clustering dengan struktur clustering yang lain yang didapatkan dari metode clustering yang sama tetapi nilai parameternya dimodifikasi.

Validasi cluster yang peneliti gunakan adalah validasi internal, karena Validasi ini didasarkan pada evaluasi hasil clustering dalam konsep kuantitatif yang mana data banjir ini juga berupa data kuantitatif. Terdapat beberapa metode atau cara pada validasi internal, diantaranya connectivity, nilai silhouette, dan indeks Dunn.

### 2.3.1 Indeks Dunn

Indeks Dunn adalah rasio jarak terkecil antara observasi pada cluster yang berbeda dengan jarak terbesar pada masing-masing cluster data. (Irwansyah dan Faisal, 2015)

Indeks Dunn diperoleh dari hasil pembagian antara  $d_{\min}$  dengan  $d_{\max}$

$$Dunn = \frac{d_{\min}}{d_{\max}}$$

dengan

$d_{\min}$  = jarak terkecil antara observasi pada *cluster* yang berbeda

$d_{\max}$  = jarak terbesar pada masing-masing *cluster* data

### 2.3.2 Indeks Silhouette

Indeks Silhouette dihitung sebagai derajat kepercayaan dalam proses clustering pada suatu pengamatan dengan cluster yang dikatakan terbentuk baik bila nilai indeks mendekati 1 dan kondisi sebaliknya jika nilai indeks mendekati angka -1 (Irwansyah dan Faisal, 2015).

$$S_{(i)} = \frac{b_{(i)} - a_{(i)}}{\max(a_{(i)}, b_{(i)})}$$

dengan

$a_{(i)}$  = jarak rata-rata antara  $i$  dan seluruh pengamatan lainnya

$b_{(i)}$  = jarak rata-rata antara  $i$  dengan pengamatan pada cluster terdekat

### 2.3.3 Indeks Connectivity

$$Conn(C) = \sum_{i=1}^N \sum_{j=1}^L X_{i,nni(j)}$$

Dengan

$nn_{i(j)}$  = pengamatan tetangga terdekat (*nearest neighbor*)  $i$  ke  $j$  dan  $L$

$nn_{i(j)}$  = sebagai parameter yang menentukan jumlah tetangga yang berkontribusi pada pengukuran *connectivity*

## 2.4 Uji Asumsi Analisis Cluster

Dalam analisis cluster ada dua asumsi yang harus dilakukan sebelum proses analisis cluster dilakukan

### 2.4.1 Uji Korelasi

Sebelum melakukan proses analisis cluster terdapat asumsi yang harus dipenuhi yaitu tidak terjadi hubungan antar variabel dan tidak terjadi korelasi.

Untuk mengetahui tidak terjadinya korelasi, maka dilakukan uji korelasi menggunakan uji koefisien korelasi *Pearson*. Koefisien korelasi diartikan sebagai tolak ukur hubungan linier variabel acak  $X_1$  dan  $X_2$  yang dilambangkan dengan  $r$  (Bluman, 2004).

$$r = \frac{n \sum_{i=1}^n x_{i1}x_{i2} - (\sum_{i=1}^n x_{i1})(\sum_{i=1}^n x_{i2})}{\sqrt{[n \sum_{i=1}^n x_{i1}^2 - (\sum_{i=1}^n x_{i1})^2][n \sum_{i=1}^n x_{i2}^2 - (\sum_{i=1}^n x_{i2})^2]}} \quad (2.13)$$

Dengan:

$r$  = koefisien korelasi pearson

$n$  = banyaknya data

$x_{i1}$  = nilai data ke-  $i$  pada variabel  $x_1$

$x_{i2}$  = nilai data ke-  $i$  pada variabel  $x_2$

Untuk mengetahui signifikansi koefisien korelasi pearson ( $r$ ) maka digunakan hipotesis yaitu:

$H_0: P = 0$  lawan  $H_1: P \neq 0$

Pengujian signifikansi tersebut dibandingkan dengan r-tabel. Dengan ketentuan apabila  $r_{hitung} \geq r_{tabel(a;n-2)}$  maka  $H_0$  ditolak dan apabila  $r_{hitung} < r_{tabel(a;n-2)}$  maka  $H_0$  diterima.

## 2.5 Covid-19

*Coronavirus disease* 2019 (COVID-19) adalah penyakit menular yang disebabkan oleh virus corona yang baru ditemukan. Coronavirus merupakan keluarga besar virus yang menyebabkan penyakit pada manusia dan hewan. *Coronavirus disease* 2019 (COVID-19) ini disebabkan *coronavirus*, yaitu virus yang menginfeksi pada sistem pernafasan. COVID-19 pertama kali terjadi di Wuhan Cina pada akhir Desember 2019, dan menyebar secara cepat hampir

seluruh negara di dunia salah satunya Indonesia. *World Health Organization* (WHO) menetapkan sebagai pandemi pada tanggal 12 Maret 2020. Pandemi ini hampir menyebar ke penjuru dunia dalam waktu yang bersamaan. Jumlah penyebaran virus ini bertambah signifikan dan berkelanjutan secara global (Widyaningrum, 2020).

Gejala umum dari Covid-19 ini yaitu demam (suhu > 38°C), batuk kering, kelelahan, dan sesak nafas. Sebagian pasien juga mengalami pegal dan nyeri, hidung tersumbat. Jika seseorang dalam 14 muncul gejala tersebut dan pernah melakukan perjalanan disuatu daerah yang terjangkit, atau pernah kontak langsung dengan penderita Covid-19, maka orang tersebut akan dilakukan pemeriksaan untuk memastikan diagnosisnya. Gejala lain dari Covid-19 ini yaitu pilek, sakit tenggorokan, batuk, dan demam. Coronavirus ini dapat menyebar melalui droplet dari hidung atau mulut pada saat batuk atau bersin. Droplet tersebut jatuh pada benda sekitar. Kemudian jika seseorang menyentuh benda yang sudah terkontaminasi dengan droplet tersebut, lalu orang itu menyentuh mata, hidung, mulut maka orang tersebut bisa terinfeksi Covid-19. Coronavirus ini tidak memiliki batasan umur, namun orang tua lebih rentan terkena, dan orang yang memiliki masalah kesehatan seperti tekanan darah tinggi, gangguan jantung, atau diabetes.

Istilah dalam Covid-19 antara lain:

### **1. Kasus Probable**

Probable yaitu orang yang masuk dalam kategori suspek dan memiliki gejala ISPA, gagal nafas, atau meninggal dunia, namun belum ada hasil tes yang menunjukkan bahwa dirinya positif Covid-19.

## 2. Kasus Konfirmasi

Kasus konfirmasi Covid-19 yaitu orang yang sudah dinyatakan terkena Covid-19 berdasarkan hasil pemeriksaan laboratorium. Kasus ini bisa terjadi pada orang yang mempunyai gejala ataupun tidak mempunyai gejala.

## 3. Kasus Meninggal Dunia

Kasus Meninggal yaitu kondisi ketika orang yang termasuk dalam kategori kasus probable dan sudah di konfirmasi terkena Covid-19.

## 4. Aktif

Kondisi ketika seseorang melakukan kontak dengan orang yang termasuk kedalam kategori konfirmasi dan probable.

### 2.6 Kajian Integrasi topik dalam Islam

Hampir seluruh aspek yang ada didunia ini ada di dalam Al-Qur'an, seperti pengelompokan (klaster) dan Covid-19 ini. Meskipun tidak di jelaskan secara terang-terangan, jadi masih perlu penafsiran untuk memahami ayat tersebut secara utuh.

#### 2.6.1 Kajian Integrasi Islam tentang wabah

Jauh sebelum kasus Covid-19 ini terjadi, pada zaman Rasulullah SAW sudah terdapat sebuah wabah yang dikenal dengan istilah tho'un, penyakit tho'un ini juga ada pada masa Kholifah Umar bin Khattab.

Di dalam Islam sudah diajarkan bagaimana cara kita menghadapi saat wabah terjadi. Seperti hadits yang menjelaskan tentang adanya tho'un.

عَنْ عَامِرِ بْنِ سَعْدِ بْنِ أَبِي وَقَّاصٍ، عَنِ أَبِيهِ، أَنَّهُ سَمِعَهُ يُسْأَلُ أُسَامَةَ بْنَ زَيْدٍ مَاذَا سَمِعْتَ مِنْ رَسُولِ اللَّهِ ﷺ فِي الطَّاعُونِ؟ فَقَالَ أُسَامَةُ: قَالَ رَسُولُ اللَّهِ ﷺ الطَّاعُونُ رَجْسٌ أُرْسِلَ عَلَى طَائِفَةٍ مِنْ بَنِي إِسْرَائِيلَ، أَوْ عَلَى مَنْ

كَانَ قَبْلَكُمْ، فَإِذَا سَمِعْتُمْ بِهِ بِأَرْضٍ، فَلَا تَقْدَمُوا عَلَيْهِ، وَإِذَا وَقَعَ بِأَرْضٍ، وَأَنْتُمْ بِمَا فَلَا تَخْرُجُوا فِرَارًا مِنْهُ». قَالَ أَبُو النَّضْرِ: «لَا يُخْرِجُكُمْ إِلَّا فِرَارًا مِنْهُ». مُتَّفَقٌ عَلَيْهِ .

*“Dari Amir bin Sa'ad bin Abi Waqqash dari ayahnya, bahwa dia mendengar ayahnya bertanya kepada Usamah bin Zaid: Apa yang kau dengar dari Rasulullah tentang wabah thaun? Usamah berkata: Rasulullah shallallahu 'alaihi wasallam bersabda: "Wabah thaun adalah adzab (siksaan) yang dikirim untuk sekelompok orang dari bani Israil, atau untuk orang sebelum kalian, jika kalian mendengar penyebaran wabah pada suatu wilayah maka jangan kalian berani mendatanginya, sedang jika wabah menyebar pada daerah kalian maka janganlah kalian keluar untuk menghindarinya". Abu An-Nadlar berkata: "Janganlah kalian keluar darinya melainkan untuk menyelamatkan diri".*

Berdasarkan hadits tersebut salah satu cara apabila terjadi suatu wabah yang melanda suatu daerah adalah jangan memasuki daerah tersebut dan apabila suatu wabah terjadi maka tidak di perbolehkan keluar dari daerah tersebut. Kedua cara tersebut adalah salah satu langkah kita mengurangi rantai penyebaran Covid-19.

Di dalam hadits lain juga menjelaskan bahwa:

حَدَّثَنَا يَحْيَى بْنُ سَعِيدٍ عَنْ شُعْبَةَ قَالَ حَدَّثَنِي أَبُو بَكْرِ بْنُ حَفْصٍ عَنْ ابْنِ الْمُصْبِحِ أَوْ أَبِي الْمُصْبِحِ عَنْ ابْنِ السَّمْطِ عَنْ عَبْدِ اللَّهِ بْنِ الصَّامِتِ قَالَ قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ عَبْدَ اللَّهِ بْنُ رَوَاحَةَ فَمَا تَحَوَّزَ لَهُ عَنْ فِرَاشِهِ فَقَالَ مَنْ شُهِدَاءُ أُمَّتِي قَالُوا قَتَلُ الْمُسْلِمِ شَهَادَةٌ قَالَ إِنَّ شُهِدَاءَ أُمَّتِي إِذَا لَقِيلُ قَتَلُ الْمُسْلِمِ شَهَادَةٌ وَالطَّاعُونَ شَهَادَةٌ وَالْبَطْنُ وَالْعَرَقُ وَالْمَرْأَةُ يُقْتَلُهَا وَلَدُهَا جَمَعَاءَ

*“Telah menceritakan kepada kami [Yahya bin Sa'id] dari [Syu'bah] berkata; Telah bercerita kepadaku [Abu Bakar bin Hafsh] dari [Ibnu Al Mushabbih atau Abu Al Mushabbih] dari [Ibnu As Simth] dari [Ubadah bin Ash Shamit] berkata: Rasulullah Shallallahu'alaihiwasallam menjenguk 'Abdullah bin Rowahah, ketika beliau Shallallahu'alaihiwasallam duduk di ranjangnya, beliau Shallallahu'alaihiwasallam bersabda: "Siapakah orang yang syahid dari umatku?" mereka berkata: orang muslim yang terbunuhlah yang terhitung syahid. Rasulullah Shallallahu'alaihiwasallam bersabda: "Kalau begitu orang yang syahid dari umatku sangat sedikit, padahal terbunuhnya seorang muslim terhitung syahid, kematian karena wabah thaun terhitung syahid, kematian karena sakit perut terhitung syahid, kematian karena tenggelam terhitung syahid dan seorang wanita yang mati karena melahirkan anaknya terhitung syahid.".*

Hadits tersebut menjelaskan bahwa orang yang mendapat pahala mati syahid salah satunya adalah orang yang terkena wabah atau tha'un. Penyakit

wabah ini sama dengan tha'un, apabila terjadi suatu wabah disuatu tempat maka tidak boleh berada di tempat tersebut dan orang yang terkena wabah tidak boleh keluar dari daerah tersebut. Seseorang yang berada di tempat wabah lalu meninggal dunia dia mendapat pahala syahid karena dari kesabarannya menghadapi penyakit yang sangat sulit di hindari.

### 2.6.2 Kajian Integrasi Islam tentang *Cluster*

*Cluster* atau kelompok di dalam Islam banyak menjelaskan tentang hal-hal yang berkaitan dengan pengelompokan. Seperti salah satu ayat dalam Al-Qur'an:

إِنَّمَا الْمُؤْمِنُونَ الَّذِينَ إِذَا ذُكِرَ اللَّهُ وَجِلَتْ قُلُوبُهُمْ وَإِذَا تُلِيَتْ عَلَيْهِمْ آيَاتُهُ زَادَتْهُمْ إِيمَانًا وَعَلَىٰ رَبِّهِمْ يَتَوَكَّلُونَ (٢)

الَّذِينَ يُقِيمُونَ الصَّلَاةَ وَمِمَّا رَزَقْنَاهُمْ يُنْفِقُونَ (٣) أُولَٰئِكَ هُمُ الْمُؤْمِنُونَ حَقًّا ۗ لَهُمْ دَرَجَاتٌ عِنْدَ رَبِّهِمْ

وَمَغْفِرَةٌ وَرِزْقٌ كَرِيمٌ (٤)

*“Sesungguhnya orang-orang yang beriman ialah mereka yang disebut nama Allah bergetarlah hati mereka, dan apabila di bacakan ayat-ayat Nya bertambah iman mereka (karenanya, dan hanya kepada Tuhanlah mereka bertawakal, yaitu orang-orang yang mendirikan sholat dan yang menafkahkan sebagian rizki yang Kami berikan kepada mereka. Itulah orang-orang yang beriman dengan sebenar-benarnya. Maka akan memperoleh beberapa derajat ketinggian di sisi Tuhannya dan ampunan serta rizki yang mulia” (QS. Al Anfal/8:2-4).*

Menurut Ibnu Katsir penjelasan Surat Al-Anfal ayat 2-4 ini adalah orang yang munafik hatinya tidak akan sedikitpun mengingat Allah dalam menjalankan kewajibannya. Mereka juga tidak beriman kepada ayat-ayat Allah, tidak bertawakal, tidak menjalankan sholat saat sendirian, tidak mengeluarkan zakat dalam harta kekayaan mereka. Maka Allah memberitahukan bahwa mereka bukanlah orang-orang yang beriman. Kemudian Allah mensifati orang-

orang beriman apabila disebut nama Allah, bergetarlah hatinya, mereka melaksanakan kewajiban-kewajiban dan apabila dibacakan ayat-ayat Allah bertambah keimanannya. Mendirikan sholat artinya menjaga waktu penunaianya, menyempurnakan setiap gerakan sholat termasuk rukuk, sujud, membaca tasyahud dan sholawat, sifat lain dari orang beriman adalah menafkahkan sebagian rizki termasuk mengeluarkan zakat kepada yang berhak menerimanya karena sejatinya rizki ini adalah tiipan dari Allah SWT. Mereka yang memperoleh derajat ketinggian di sisi Tuhannya artinya adalah mereka akan mendapatkan tempat dan kedudukan derajatnya lebih tinggi daripada sebagian orang yang berada dibawahnya, orang yang berada di tingkatan tertinggi dapat melihat orang yang kedudukannya dibawahnya, akan tetapi orang yang berada dibawahnya tidak mempunyai pandangan bahwa tidak seorang pun lebih utama dari pada dirinya. (Abu Al-Fida' Ibnu Katsir, 1999).

Surat Al-Anfal ayat 2-4 tersebut menjelaskan tentang orang-orang yang beriman dan bertawakal kepada Allah SWT. Apabila seseorang memiliki salah satu dari ciri-ciri yang sama dalam ayat tersebut maka orang tersebut termasuk dalam kelompoknya. Begitu pula *cluster* , jika suatu objek memiliki karakteristik yang sama maka objek tersebut termasuk *cluster* yang sama.



## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Pendekatan Penelitian**

Pada penelitian ini menggunakan pendekatan studi literatur dan deskriptif kuantitatif. Yang mana studi literatur yang digunakan berupa pengumpulan bahan pustaka yang diperoleh dari buku dan jurnal. Sedangkan pendekatan deskriptif kuantitatif yaitu dengan menganalisis data sesuai dengan kebutuhan peneliti.

#### **3.2 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data mengenai kasus Covid-19 di Provinsi Jawa Timur pada tanggal 29 Mei 2023. Data diperoleh dari website resmi Pemerintah Provinsi Jawa Timur yaitu *infocovid19.jatimprov.go.id*.

#### **3.3 Variabel Penelitian**

Pada penelitian ini variabel yang digunakan yaitu:

$X_1$ : Kasus Probable

$X_2$ : Kasus Konfirmasi

$X_3$ : Kasus Meninggal Dunia

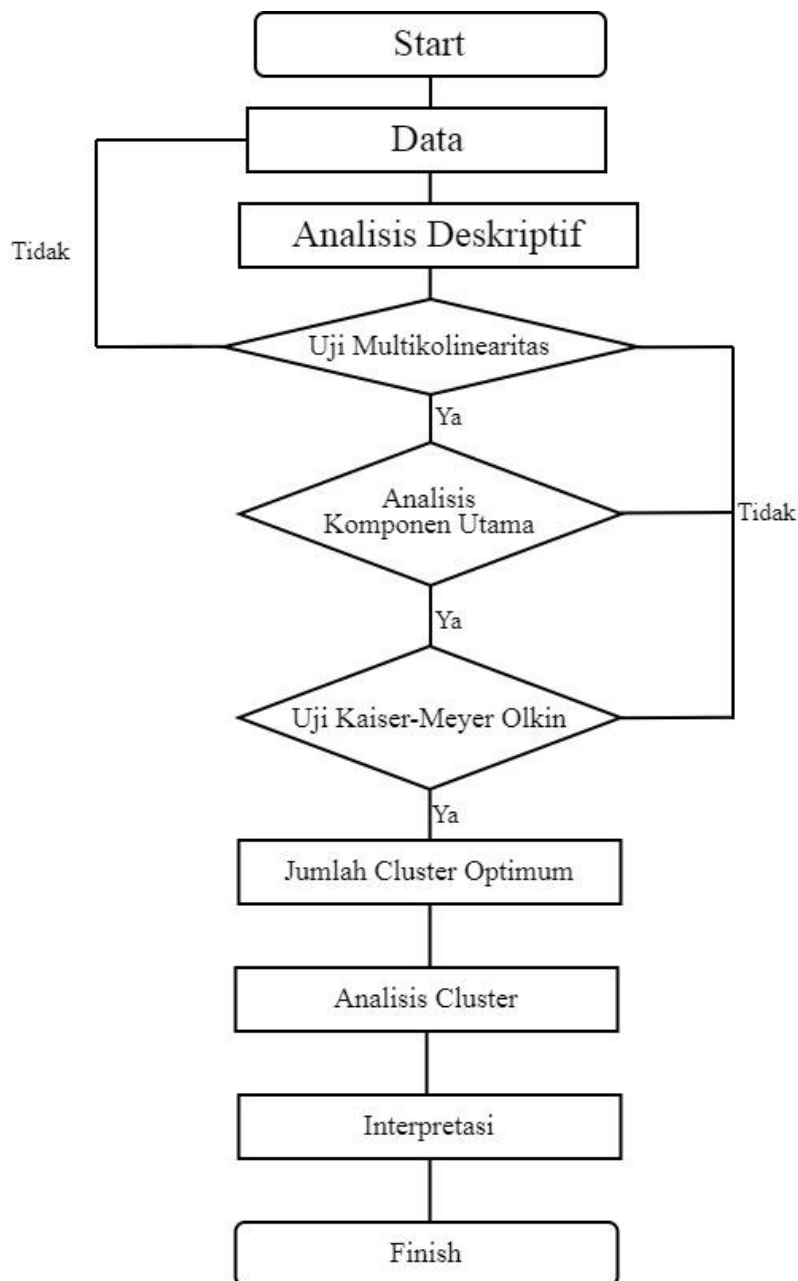
$X_4$ : Kasus Aktif

#### **3.4 Metode Analisis Data**

Langkah-langkah yang dilakukan dalam penelitian ini adalah:

1. Mengumpulkan data yang akan dianalisis
2. Melakukan uji korelasi dengan uji koefisien korelasi *Pearson*
3. Menghitung jarak menggunakan Jarak Pearson Correlation

4. Membentuk *cluster* dengan metode ward
5. Menentukan jumlah *cluster* dengan anggota *cluster* yang sudah terbentuk
6. Membentuk *cluster* yang akan dijadikan karakteristik dari *cluster* tersebut
7. Menarik Kesimpulan



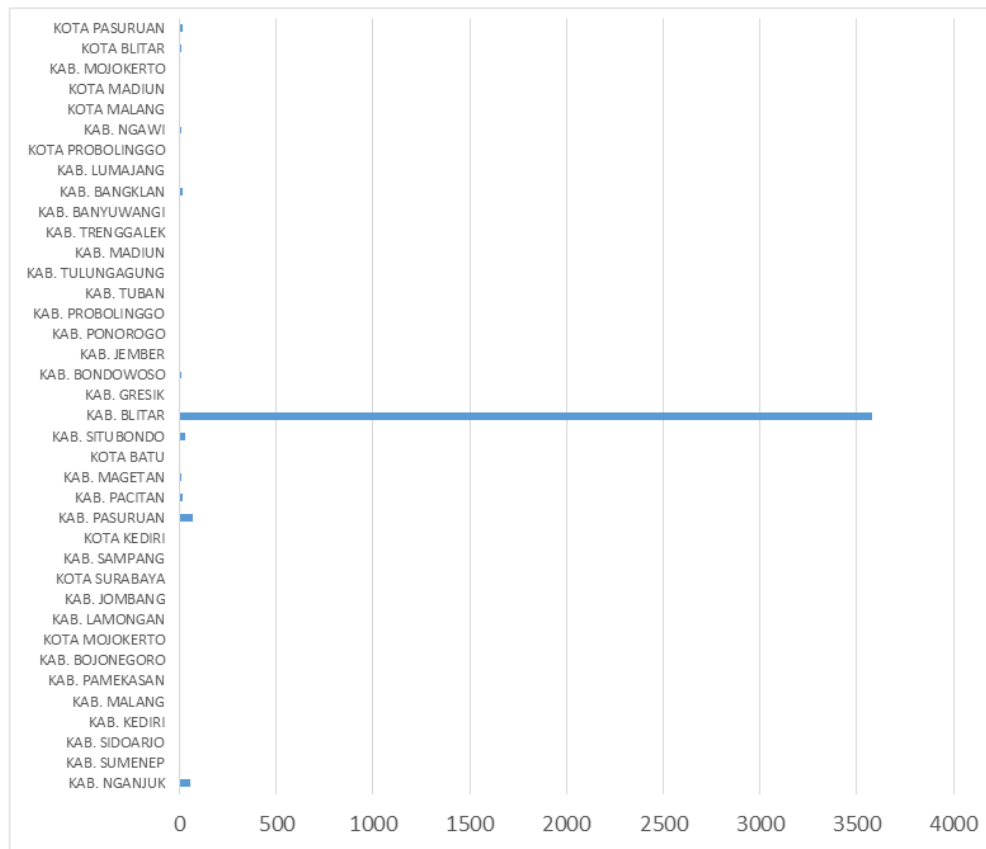
Gambar 3. 1 *Flowchart* Penelitian

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Analisis Deskriptif

Tujuan dari analisis deskriptif adalah untuk mempermudah pembaca dalam memahami suatu gambaran dari sebuah data. Pada penelitian ini dilakukan analisis deskriptif mengenai data diagnosis mengenai kasus covid-19 di Kabupaten/Kota di Jawa Timur.



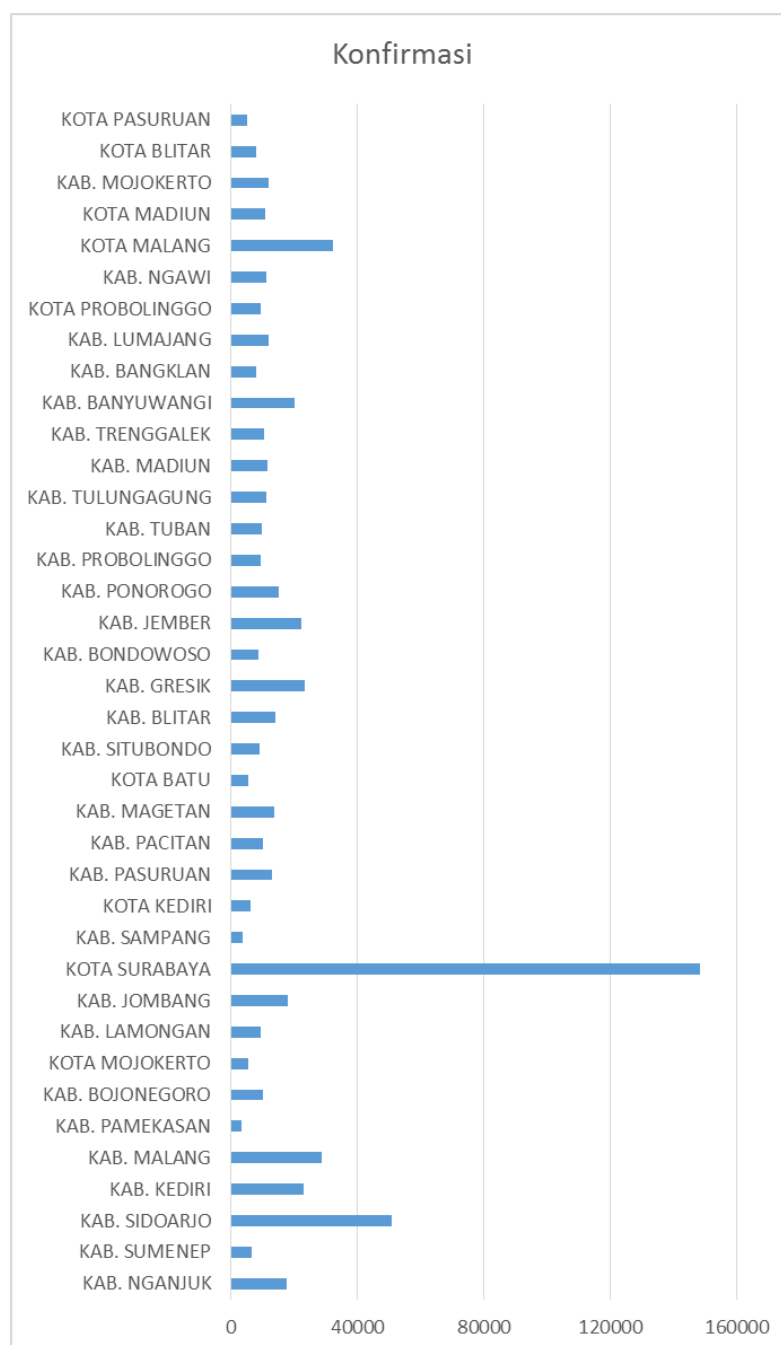
Gambar 4. 2 Kasus Probable Covid-19 Kabupaten/Kota di Jawa Timur

##### a. Probable

Gambar 4.1 dapat dilihat bahwa terdapat daerah dengan kategori suspek dan memiliki gejala ISPA, gagal napas atau meninggal dunia namun belum ada hasil tes yang menunjukkan bahwa dirinya positif covid-19 yaitu pada kab. Blitar dengan jumlah 3581. Selain itu juga terdapat wilayah yang sedikit yaitu Kab. Ngajuk, Sumenep, Sidoarjo, Kediri, Malang Pamekasan, Bojonegoro, Mojokerto, Lamongan, Jombang, Sampang, Gresik, Jember, Ponorogo,

Probolinggo, Tuban, Tulungagung, Madiun, Trenggalek, Banyuwangi, Lumajang. Dalam definisi operasional kasus covid-19, kasus probable termasuk salah satu dari delapan kategori kasus covid-19, bersama dengan kasus suspek dan konfirmasi. Dalam hal ini, diagnosis dini sangat penting dalam pencegahan COVID-19 dan memungkinkan implementasi tindakan untuk membatasi penyebaran melalui identifikasi.

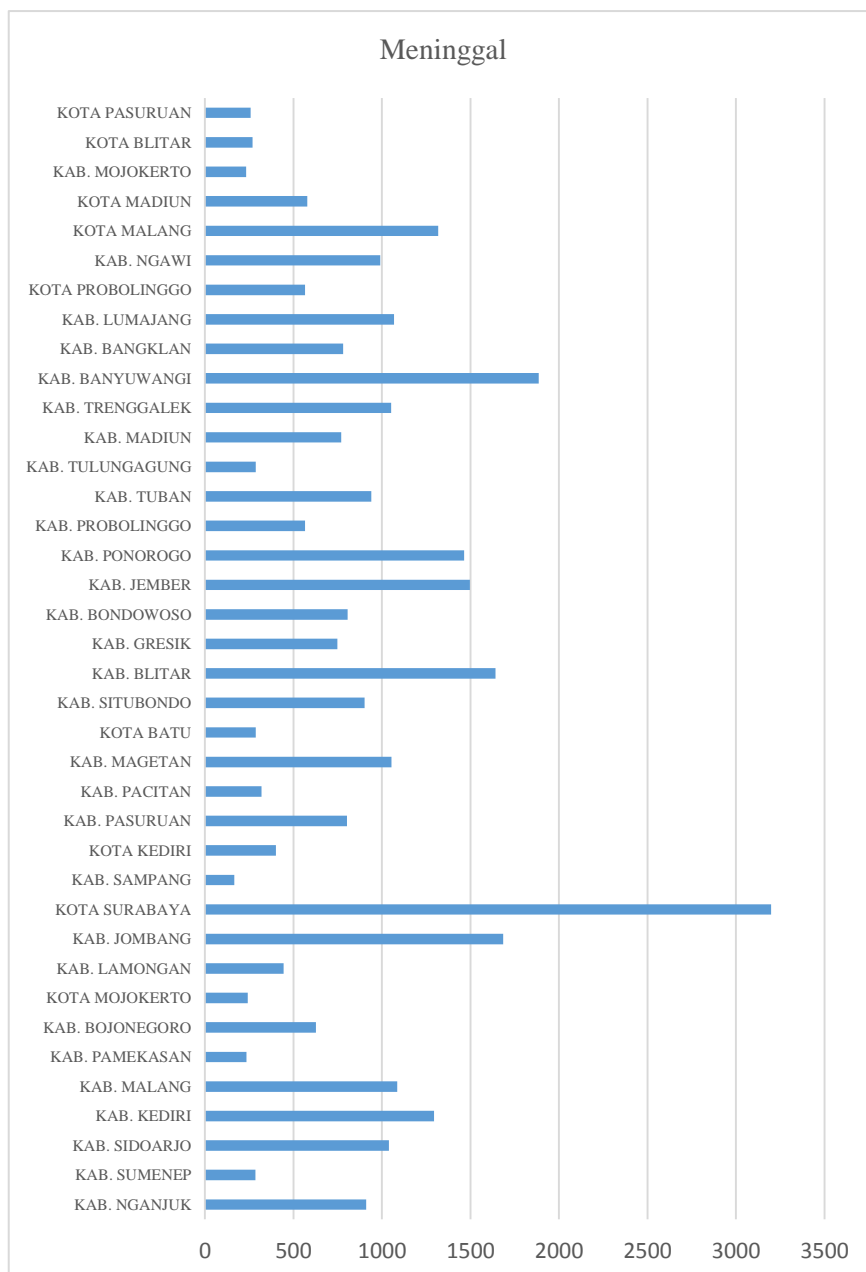
## b. Konfirmasi



Gambar 4.3 Kasus Konfirmasi Covid-19 Kabupaten/Kota di Jawa Timur

Gambar 4.2 menunjukkan jumlah kasus konfirmasi covid-19 dimana orang yang sudah dinyatakan terkena covid-19 berdasarkan hasil pemeriksaan laboratorium. Kasus ini bias terjadi pada orang yang mempunyai gejala ataupun tidak mempunyai gejala, dimana kejadian terbanyak pada Kabupaten/Kota di Jawa timur terdapat pada daerah Kota Surabaya sebanyak 148.587. Selain itu juga terdapat Kabupaten/Kota di Jawa Timur dengan kejadian paling sedikit yaitu Kabupaten Pamekasan sebanyak 3.348. Kontak erat harus melakukan tindakan pencegahan seperti mencuci tangan, memakai masker dan menjaga jarak dengan orang lain untuk mencegah penebaran virus.

### c. Meninggal

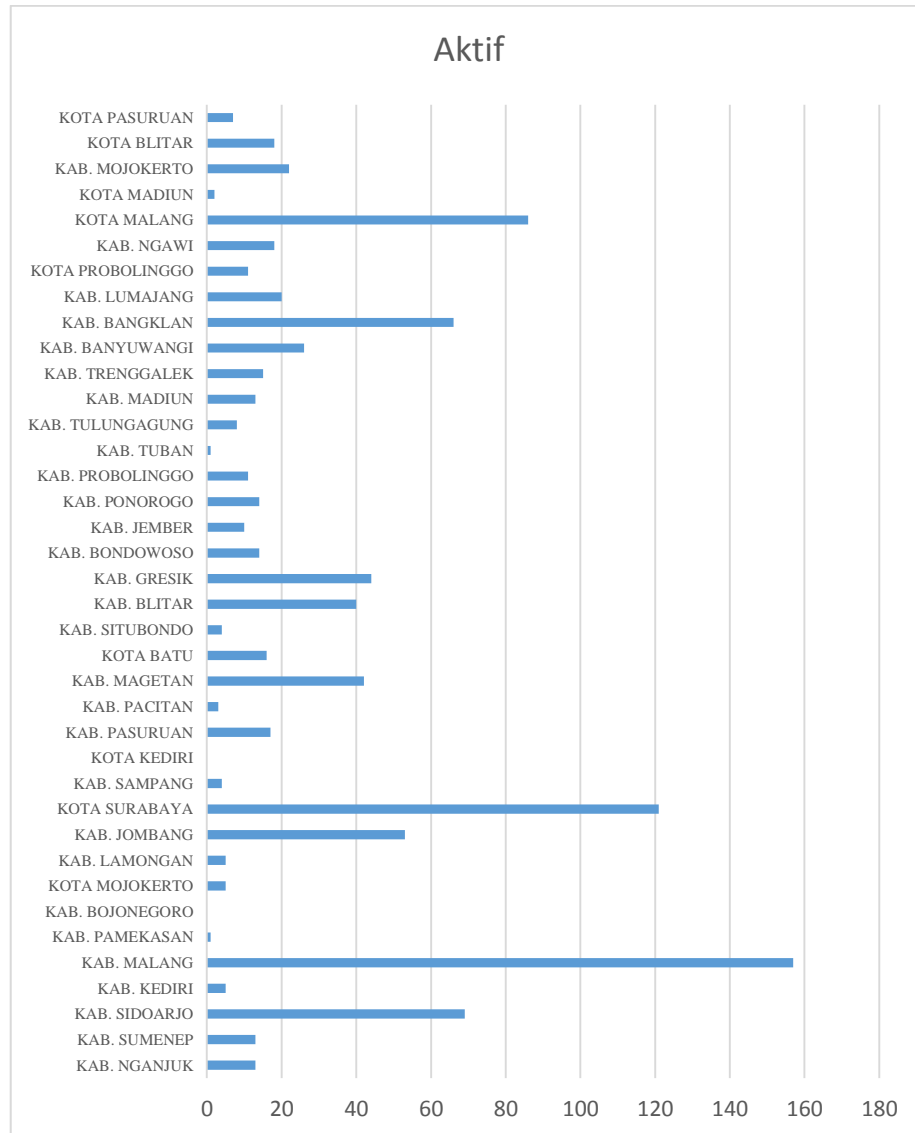


**Gambar 4.3 Kasus meninggal Covid-19 Kabupaten/Kota di Jawa Timur**

Dari gambar 4.3 menunjukkan angka kematian covid-19 pada Kabupaten/Kota di Jawa Timur dimana kejadian terbanyak terjadi pada daerah Kota Surabaya dengan kejadian sebanyak 3.198. selain itu juga wilayah pada dengan angka kematian akibat covid-19 rendah yaitu Kabupaten Sampang dengan kejadian sebanyak 166. Penyebaran covid dan pengurangan angka kematian akibat covid semakin hari semakin naik. Dengan mengikuti beberapa aturan protocol kesehatan yang telah ditetapkan oleh pemerintah dapat

mengurangi penyebaran salah satunya dengan menghindari kerumunan dan mengurangi mobilitas.

#### d. Aktif



**Gambar 4. 4 Kasus aktif Covid-19 Kabupaten/Kota di Jawa Timur**

Dari gambar 4.4 diatas menunjukkan angka aktif covid-19 pada Kabupaten /Kota di Jawa Timur dimana kejadian terbanyak di Jawa Timur dengan kejadian terbanyak terdapat pada daerah Kabupaten Malang sebanyak 157, Kota Surabaya sebanyak 121 dan Kota Malang sebanyak 86. Daerah dengan jumlah angka covid-19 paling sedikit yaitu Kabupaten Bojonegoro dan Kabupaten

Kediri dengan angka 0 pada kasusnya. Pentingnya kesadaran masyarakat tentang mematuhi protocol kesehatan dan mengikuti program vaksin covid-19 merupakan salah satu untuk mencegah penyebaran.

## 4.2 Uji Asumsi Cluster

### 4.2.1 Uji Multikolinieritas

Uji multikolinieritas merupakan terjadinya korelasi yang kuat antara dua atau lebih variabel kelompok, salah satu masalah yang harus diperhatikan dalam menganalisis multivariat, karena multikolinieritas disebabkan oleh pengaruh besar dalam menghasilkan solusi sehingga nantinya dapat mengganggu proses analisis.

**Tabel 4. 1 Matriks Korelasi**

	Probable	Konfirmasi	Meninggal	Aktif
Probable	1.00	-0.23	0.21	0.06
Konfirmasi	-0.02	1.00	0.75	0.64
Meninggal	0.21	0.75	1.00	0.56
Aktif	0.06	0.64	0.5	1.00

Menurut Gujarati (1995) bila korelasi antara dua variabel bebas melebihi 0.8 maka multikolinieritas menjadi masalah yang serius. Apabila korelasi antara variabel penjelas tidak lebih besar dibandingkan korelasi variabel terikat dengan masing-masing variabel penjelas, maka dapat dikatakan tidak terdapat masalah serius. Pada tabel 4.1 diketahui bahwa nilai korelasi  $< 0.8$  maka hal itu menandakan tidak terdapat masalah serius.



### 2.4.2 Principal Component Analysis (PCA)

*Principal Component Analysis* (PCA) atau analisis komponen utama merupakan solusi jika dalam proses analisis clustering terjadi multikolinieritas. Tujuan dari analisis PCA adalah untuk mereduksi variabel yang ada menjadi lebih sedikit tanpa harus kehilangan informasi yang termuat dari data asli. Dalam menentukan jumlah komponen utama yang dihasilkan, dapat dilihat dari *eigenvalue* yang menunjukkan lebih dari satu. Hasil PCA seperti pada Tabel 4.2.

**Tabel 4. 2 Tabel *Eigenvalues***

Komponen	Total	% of variance	Cumulative %
1	1.524	0.581	0.581
2	1.0109	0.2555	0.8365
3	0.6764	0.1144	0.9509
4	0.4433	0.0491	1.0000

Dari tabel 4.2 menunjukkan nilai *eigenvalue* dan proporsi kumulatif untuk tiap faktornya. Dari hasil tersebut, komponen 1 dan 2 memiliki nilai *eigenvalue* lebih dari satu. Sehingga dipilihlah komponen yang terbentuk dari analisis PCA ini sebanyak 2 komponen utama. Komponen pertama memiliki *eigenvalue* sebesar 1.524 dengan varian sebesar 0.581%. komponen kedua memiliki nilai *eigenvalue* sebesar 1.0109 dengan varian sebesar 0.2555%. Dengan menggunakan 2 komponen utama, kedua komponen tersebut telah mampu menjelaskan 0.8365% keragaman data. Dibawah ini merupakan hasil dari PCA.

Tabel 4. 3 Hasil PCA

PC1	PC2
-0.141	-0.037
-1.038	-0.146
1.689	-0.543
0.221	-0.079
2.564	-0.599
-1.356	-0.093
-0.819	-0.078
-1.233	-0.121
-0.933	-0.12
1.231	-0.105
7.037	-1.121

#### 4.2.3 Uji Kaiser-Meyer Olkin

Uji KMO digunakan untuk mengukur kecukupan pengambilan sampel untuk setiap indicator. Apabila nilai KMO lebih dari 0,5 maka asumsi sampel mewakili populasi terpenuhi atau sampel representatif. Hasil nilai KMO dapat dilihat pada tabel

Tabel 4. 4 Koefisien Korelasi

<i>Kaiser Mayer Olkin</i> (KMO)	0.62
------------------------------------	------

Berdasarkan hasil nilai KMO pada tabel 4.4 didapatkan nilai KMO sebesar 0.62 yang artinya nilai ini melebihi 0.5 yang berarti bahwa sampel yang mewakili populasi atau sampel representative sehingga analisis bias dilanjutkan ke tahap selanjutnya.

#### 4.3 Penentuan jumlah *Cluster* Optimum

Pada penelitian ini menggunakan tiga pendekatan yaitu indeks *connectivity*, indeks *dunn*, dan *silhouette*. Pengambilan jumlah *cluster* dengan melihat nilai indeks *connectivity* paling kecil, nilai indeks *dunn* paling besar, dan nilai

silhouette mendekati 1. Berikut merupakan hasil yang didapatkan dari validasi cluster *hierarchical* seperti pada tabel

**Tabel 4.5 Penentuan Jumlah Cluster Optimum**

Metode Cluster	Indeks	Jumlah Cluster			
		2	3	4	5
<i>Hierarchical Clustering</i>	<i>Connectivity</i>	2.929	5.857	11.294	13.877
	Dunn	0.669	1.137	0.312	0.312
	<i>Silhouette</i>	0.788	0.782	0.623	0.553

Berdasarkan pada hasil tiga pendekatan di tabel 4.5, diperoleh bahwa pada algoritma *hierarchical clustering* dilihat dari nilai indeks *connectivity* yang paling kecil dengan nilai 2.929 pada *cluster* 2, nilai dunn dengan nilai paling besar yaitu 1.137 pada *cluster* 3, dan nilai *silhouette* yang mendekati 1 yaitu 0.788 pada *cluster* 2. Sehingga, pada algoritma *hierarchical clustering* diperoleh *cluster* yang paling baik sebanyak 2. Maka pada penelitian ini menggunakan jumlah *cluster* sebanyak 2.

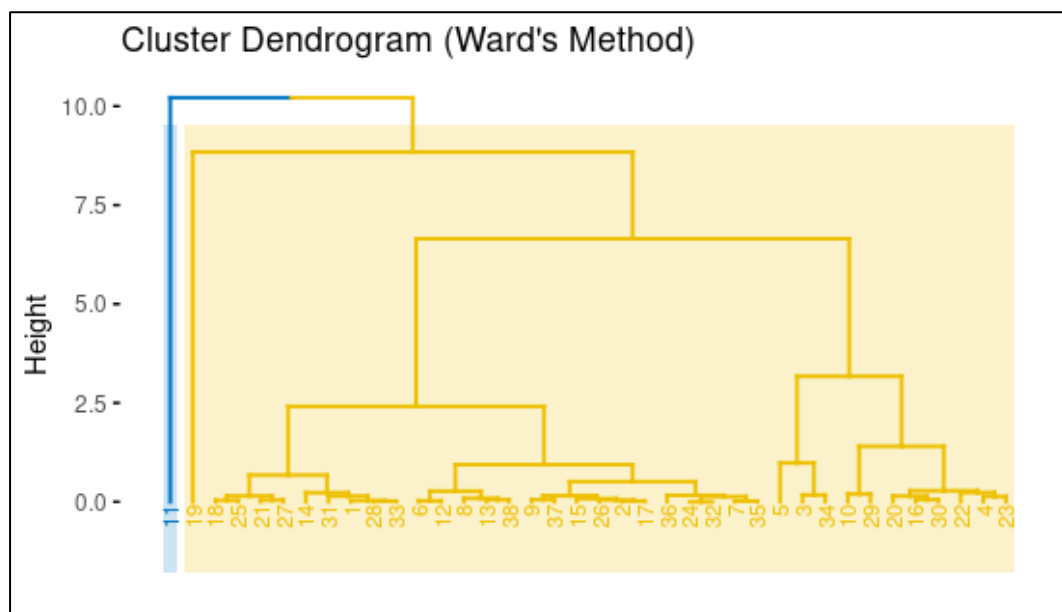
#### 4.4 Clustering Hierarchical

*Clustering Hierarchical* adalah suatu metode analisis cluster yang dilakukan secara bertahap dan bertingkat sehingga membentuk tingkatan seperti pohon. Dalam menentukan jumlah *cluster*, metode ini menghasilkan urutan partisi dengan menggabungkan atau membagi *cluster*. Hasil dari metode ini dapat disajikan dengan dendrogram. Adapun 4 metode *agglomerative* dalam pembagian *cluster* yaitu *single linkage*, *average linkage*, *complete linkage*, dan *ward*. Penentuan metode terbaik pada *cluster hierarchical* dapat ditentukan dengan *agglomerative coefficient*, pada tabel 4.6 merupakan hasil dari *agglomerative coefficient*.

**Tabel 4.6 Agglomerative Coefficient**

Metode	Agglomerative coefficient
<i>single linkage</i>	0.9394357
<i>average linkage</i>	0.9371964
<i>complete linkage</i>	0.9407347
<i>Ward</i>	0.9411024

Berdasarkan tabel 4.6 didapatkan bahwa metode *hierarchical* yang menghasilkan *cluster* terbaik yaitu metode *Ward* karena nilai *agglomerative coefficient* lebih besar dibandingkan dengan metode *hierarchical* lainnya. Dendrogram dengan metode *ward* seperti gambar 4.5.

**Gambar 4.5 Dendrogram Ward**

Dendrogram yang diperoleh dari metode *ward* terlihat terpartisi dengan sangat baik. Pada setiap cluster, terlihat nilai *height* yang rendah, hal ini dikarenakan cara kerja metode *ward* yaitu meminimumkan nilai *within sum of*

*squared* (wss) tiap *cluster*. Pada dendrogram diatas terdapat 1 kabupaten/kota pada *cluster* 1, dan terdapat 37 kabupaten/kota pada *cluster* 2.

Untuk mengetahui anggota kelompok yang terbentuk pada masing-masing *cluster* dari metode *ward*, seperti pada tabel berikut.

**Tabel 4. 7 Anggota Cluster**

Cluster	Jumlah	Kabupaten/Kota
1	1	Kota Surabaya
2	37	Kabupaten Nganjuk, Kabupaten Sumenep, Kabupaten Sidoarjo, Kabupaten Kediri, Kabupaten Malang, Kabupaten Pamekasan, Kabupaten Bojonegoro, Kabupaten Mojokerto, Kabupaten Lamongan, Kabupaten Jombang, Kabupaten Sampang, Kota Kediri, Kabupaten Pasuruan, Kabupaten Pacitan, Kabupaten Magetan, Kota Batu, Kabupaten Situbondo, Kabupaten Blitar, Kabupaten Gresik, Kabupaten Bondowoso, Kabupaten Jember, Kabupaten Ponorogo, Kabupaten Probolinggo, Kabupaten Tuban, Kabupaten Tulungagung, Kabupaten Madiun, Kabupaten Trenggalek, Kabupaten Banyuwangi, Kabupaten Bangkalan, Kabupaten Lumajang, Kota Probolinggo,

		Kabupaten Ngawi, Kota Malang, Kota Madiun, , Kabupaten Mojokerto, Kota Blitar, Kota Pasuruan.
--	--	---

#### 4.5 Interpretasi Metode Terbaik

Berdasarkan hasil perbandingan metode *cluster* terbaik menggunakan pendekatan *cluster*, didapatkan hasil yaitu hierarchial cluster menggunakan metode ward pada data mengenai kasus covid-19 berdasarkan kabupaten/kota di provinsi Jawa Timur, berikut merupakan hasil profilasi.

**Tabel 4. 8 Profilisasi Cluster Terbaik**

Cluster	Probable	Konfirmasi	Meninggal	Aktif
1	11.34	13681	830.01	23.03
2	26.67	35510	1020.5	39.50

Berdasarkan hasil tabel 4.7 didapatkan hasil profilisasi atau nilai rata-rata yang artinya dari masing-masing cluster memiliki karakteristik yang berbeda. Pada cluster 1 dan cluster terdapat perbedaan yang mencolok yaitu nilai rata-rata *cluster* 1 lebih rendah dibandingkan cluster 2. Warna hijau menunjukkan rendah, dan warna merah menunjukkan tinggi.

#### 4.6 Integrasi Al-qur'an dan Cluster

Pada Al-qur'an penjelasan terkait cluster atau pengelompokan tertuang dalam surat Al- Waqi'ah ayat 7-12 yang artinya:

وَكُنْتُمْ أَزْوَاجًا ثَلَاثَةً (٧) فَأَصْحَابُ الْمَيْمَنَةِ مَا أَصْحَابُ الْمَيْمَنَةِ (٨) وَأَصْحَابُ الْمَشْأَمَةِ مَا  
 أَصْحَابُ الْمَشْأَمَةِ (٩) وَالسَّابِقُونَ السَّابِقُونَ (١٠) أُولَئِكَ الْمُقَرَّبُونَ (١١) فِي جَنَّاتِ النَّعِيمِ  
 (١٢)

“dan kamu menjadi tiga golongan, (7) yaitu golongan kanan, alangkah mulianya golongan kanan itu, (8) dan golongan kiri, alangkah sengsaranya golongan kiri itu, (9) dan orang-orang yang paling dahulu (beriman), (10) merekalah yang paling dahulu (masuk surga).(11) Mereka itulah orang yang dekat (kepada Allah), Berada dalam surga kenikmatan, (12)”(QS.Al-Waqiah/56:7-13).

Sebagaimana penjelasan dari ayat diatas , pada hari kiamat manusia akan terbagi menjadi tiga kelompok. kelompok pertama disebut *ashab al-maimanah*, kelompok kedua dinamakan *ashab al-masy'amah*, dan kelompok ketiga disebut *as-sabiqun as-sabiqun*. Klasifikasi ini ada yang menyebutnya dengan tempat atau kedudukan manusia di hari kiamat. Seperti yang telah disebutkan pada ayat tersebut salah satunya adalah golongan kanan, manusia yang masuk ke dalam golongan kanan adalah yang memiliki sifat mulia. Begitu pula di dalam *cluster*, suatu objek akan dikelompokkan ke dalam kelompok apabila memiliki karakteristik yang sama.

Sebagaimana diinformasikan dalam redaksi ayat, pada hari kiamat manusia akan terbagi menjadi tiga kelompok. kelompok pertama disebut *ashab al-maimanah*, kelompok kedua dinamakan *ashab al-masy'amah*, dan kelompok ketiga disebut *as-sabiqun as-sabiqun*. Klasifikasi ini ada yang menyebutnya dengan tempat atau kedudukan manusia di hari kiamat, seperti penjelasan yang dikutip dalam Tafsir At-Tabari. Jadi, ada dua kelompok yang bertempat di surga dan satu kelompok di neraka. Demikian lanjutan penjelasan At-Tabari.

Penggunaan kanan dan kiri dalam surat Al-Waqiah ayat 7-12 ini tidak lain mengambil dari bagian yang sudah familiar dengan kehidupan manusia yaitu sisi kanan dan kiri. Dalam tradisi di masyarakat, menurut Ibnu Asyur sisi kanan biasa diidentikkan dengan kebaikan, kemuliaan, pertolongan dan keberkahan; sedang sisi kiri sebaliknya, diasosiasikan dengan kejelekan, kemudharatan, tidak bermanfaat dan semacamnya. Di sinilah Al-Quran berinteraksi dengan sedikit kebiasaan masyarakat Arab saat itu.

Beragam penjelasan telah disampaikan oleh para mufasir dalam mendefinisikan tiga kelompok di atas. Al-Qurthubi sendiri menyajikan banyak versi tentang identitas ashab al-maimanah, ashab al-masy'amah dan as-sabiqun. Ashab al-maimanah yaitu mereka yang dibawa di sisi kanan menuju ke surga, sedang ashab al-masy'amah adalah mereka yang dibawa di sisi kiri menuju ke neraka.

Selain itu, Al-Qurthubi juga mengutip beberapa identitas lain dari ashab al-maimanah yaitu mereka yang berada di sebelah kanan Nabi Adam yang kelak akan masuk surga. As-Samarqandi dalam Bahrul Ulum menambahkan bahwa posisi di sebelah kanan Nabi Adam itu ketika di hari kiamat; mereka juga orang-orang yang diberikan catatan amalnya dengan tangan kanan.

Ashab al-maimanah juga sebutan untuk ahl al-hasanat (orang-orang yang berbuat kebaikan) dan berkomitmen untuk selalu berbuat kebaikan. Syeikh Nawawi Al-Bantani juga menuturkan pengertian yang terakhir ini. Sementara untuk ashab al-masy'amah, singkatnya adalah kebalikan dari semua hal yang diidentikkan dengan ashab al-maimanah.



Kemudian, bagaimana dengan kelompok as-sabiqun? Jika ashab al-maimanah adalah simbol kebaikan dan ashab al-masy'amah adalah simbol kejelekan, lalu as-sabiqun simbol dari apa? Ada yang menyebut as-sabiqun dengan mereka yang pernah shalat menghadap dua kiblat (masjid al-aqsha dan ka'bah di masjid al-haram) yang berarti generasi awal Islam, lebih dulu beriman kepada Allah dan rasulNya, lebih dulu melakukan perintah-perintah Allah dan RasulNya. Kurang lebih demikian intisari penjelasan dari para mufasir tentang as-sabiqun.

Selain beberapa pengertian di atas, ada keterangan tambahan dari At-Thabari yang menarik dan sedikit berbeda dalam mendefinisikan tiga kelompok di atas. Pengertian ini dikaitkan dengan ilmu dan hawa nafsu. Mereka yang memenangkan ilmunya daripada hawa nafsunya adalah as-sabiqun, mereka yang bisa menyeimbangkan antara ilmu dan hawa nafsunya ialah ashab al-maimanah dan mereka yang ilmunya tunduk pada hawa nafsunya disebut dengan ashab al-masy'amah.

Penafsiran yang sedikit berbeda juga ditambahkan oleh Al-Qurthubi. Orang-orang yang dari awal istiqamah berbuat kebaikan hingga akhir umurnya maka ia disebut dengan as-sabiqun. Orang-orang yang berbuat dosa yang kemudian menyadari kesalahannya dan bertaubat, mereka lah ashab al-maimanah. Orang-orang yang berbuat dosa, tidak pernah mau menyadari kesalahannya dan juga tidak mau bertaubat dan memperbaiki diri, mereka itu ashab al-masy'amah.

Dari beberapa pengertian di atas, kita juga dapat mengambil petunjuk bahwa hidup bahkan mati pun tidak hanya tentang kebaikan dan kejelekan, melainkan

ada sesuatu yang lebih dari itu, di atas kebaikan yaitu kemuliaan. Tidak hanya ashab al-maimanah dan ashab al-masy'amah, tetapi ada pula as-sabiqun.

Keadaan Manusia di Akhirat Berkaitan Erat dengan Amal Perbuatannya di Dunia. Terkait dengan macam kondisi manusia dalam surat Al-Waqiah ayat 7-12, Ar-Razi dalam Mafatih Al-Ghaib mengaitkannya dengan ayat yang lain, tepatnya di surat Fatir ayat 32 yang juga menyinggung tentang tiga klasifikasi keadaan manusia. Bedanya pada ayat ini konteks ayat tidak sedang menggambarkan keadaan kiamat, melainkan tentang respon penerimaan umat Muhammad terhadap Al-Quran.

Di situ disampaikan bahwa ada tiga respon penerimaan umat Muhammad terhadap ajaran yang dibawa oleh Nabi. Pertama, dhalim li nafsih (mendzalimi dirinya sendiri), kedua, muqtasid (pertengahan), ketiga sabiq bi al-khairat (lebih dulu berbuat kebaikan). Ibn Abbas sebagaimana dikutip oleh Ibn Katsir dan juga At-Tabari menjelaskan keterkaitan tiga macam penerimaan ini dengan keadaan mereka di akhirat nanti. Dhalim li nafsih ditafsirkan dengan orang yang dzalim pada dirinya sendiri yang kelak akan diampuni (terkecuali syirik dan jika orang itu bertaubat), muqtasid yaitu mereka yang kelak mudah penghitungan amalnya sedang sabiq bi al-khairat adalah mereka yang masuk surga tanpa hisab. Dengan begitu, keadaan manusia di hari kiamat seperti yang telah dijelaskan dalam surat Al-Waqiah ayat 7-12 itu tidak tiba-tiba saja terjadi tanpa proses dan sebab.

Demikian berarti bahwa surat al-waqi'ah ayat 7-12 mengkaji tentang klasifikasi manusia di hari kiamat kelak, begitu pula di dalam *cluster*, suatu objek

akan dikelompokkan ke dalam kelompok apabila memiliki karakteristik yang sama.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Pada data mengenai kasus covid-19 berdasarkan Kabupaten/Kota di Provinsi Jawa Timur terdapat beberapa daerah yang memiliki angka kasus paling banyak. Kasus probable terbanyak pada wilayah kabupaten Blitar yaitu 3581, untuk angka kasus konfirmasi covid-19 dengan kasus tertinggi pada wilayah Kota Surabaya sebanyak 148.587 kasus. Kasus Meninggal akibat covid-19 dengan angka tertinggi pada wilayah Kota Surabaya dengan kejadian sebanyak 3.198 dan kasus aktif covid-19 terdapat pada wilayah Kabupaten Malang sebanyak 157, Kota Surabaya sebanyak 121 dan Kota Malang sebanyak 86 kasus.
2. Berdasarkan hasil cluster menggunakan algoritma *hierarchical* pada data mengenai kasus covid-19 berdasarkan Kabupaten/Kota di Provinsi Jawa Timur didapatkan terdapat 2 cluster yaitu diantaranya, cluster 1 terdiri dari 1 Kabupaten/Kota, dan cluster 2 terdiri dari 37 Kabupaten/Kota di Provinsi Jawa Timur.
3. Berdasarkan hasil perbandingan metode cluster terbaik didapatkan yaitu menggunakan algoritma *Hierarchical Clustering* metode *Ward* pada data mengenai kasus covid-19 berdasarkan Kabupaten/Kota di Provinsi Jawa Timur.

## 5.2 Saran

Meskipun nilai *agglomerative coefficient* yang dihasilkan menggunakan *Ward* pada penelitian ini relatif lebih baik dibandingkan dengan metode yang lain. Sebaiknya pada penelitian selanjutnya dapat menambahkan metode perbandingan cluster lainnya dan juga dapat dikaji konsep *clustering* lain agar didapatkan hasil yang lebih baik.

## DAFTAR PUSTAKA

- Abdullah.(2007).*Tafsir Ibnu Katsirn jilid 4*. Jakarta : Pustaka Imam Asy-Syafi'i.
- Bluman, A.(2004). *Elementary Statistics: A Step By Step Approach, 5<sup>th</sup> Edition*.  
New York: Me Graw-Hill.
- Dinas Kesehatan Jawa Timur. (2020). <http://infocovid19.jatimprov.go.id/>
- Johnson, R.A & Wichern, DW. (1992). *Applied Multivariate Statistikal Analysis Third Edition*. New Jersey: Prentice Hall International.
- Lina, R., & Eni, L. T. (2011). *Analisis Kelompok Dengan Menggunakan Metode Hierarki Untuk Pengelompokan Kabupaten/Kota Di Jawa Timur Berdasarkan Indikator Kesehatan*.
- Machfudhoh,S., & Wahyuningsih, N. (2013). Analisis Cluster Kabupaten / Kota Berdasarkan Pertumbuhan Ekonomi Jawa Timur. *Sains Dan Seni Pomits*,2 (1) ,1-8.
- Mongi ,C.E. (2015). Penggunaan Analisis Two Step Clustering untuk Data Campuran.*D'CARTESIAN*,4(1),9.<http://doi.org/10.35799/dc.4.1.2015.7251>
- Nafisah, Q., & Candra, N. E.(2017). Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur. *Zeta –Math Journal*, 3(2), 31.
- Narimawati, U. (2008) . *Metodologi Penelitian Kualitatif dan Kuantitatif, Teori dan Aplikasi*. Bnadung: Agung Media.
- Santosa, E. (2007). *Data Mining:Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Simamora B. (2005). *Analisis Multivariat Pemasaran*. PT. Gramedia Pustaka Utama: Jakarta.
- Simamora , Bilson. (2015). *Analisis Multivariat Arti dan Pemasaran*. Jakarta: Rineka Cipta
- Santoso, Singgih. (2018). *Mahir Statistik Multivariat dengan SPSS*. Jakarta :PT Elex Media Komputindo

## LAMPIRAN

### Lampiran 1

## Data Persebaran Covid-19

Kab/Kota	Kasus Covid-19				Kab/Kota	Kasus Covid-19			
	Probable	Konfirmasi	Meninggal Dunia	Aktif		Probable	Konfirmasi	Meninggal Dunia	Aktif
KAB. NGANJUK	54	17827	911	13	KAB. GRESIK	0	23465	748	44
KAB. SUMENEP	0	6627	284	13	KAB. BONDOWOSO	4	8774	805	14
KAB. SIDOARJO	0	51106	1039	69	KAB. JEMBER	0	22438	1496	10
KAB. KEDIRI	0	22989	1294	5	KAB. PONOROGO	0	15354	1465	14
KAB. MALANG	0	28829	1087	157	KAB. PROBOLINGGO	0	9353	566	11
KAB. PAMEKASAN	0	3348	235	1	KAB. TUBAN	0	9749	939	1
KAB. BOJONEGORO	0	10322	627	0	KAB. TULUNGAGUNG	0	11259	286	8
KOTA MOJOKERTO	0	5521	241	5	KAB. MADIUN	0	11821	770	13
KAB. LAMONGAN	0	9669	444	5	KAB. TRENGGALEK	0	10414	1051	15

KAB. JOMBANG	0	18013	1685	53	KAB. BANYUWANGI	0	20361	1886	26
KOTA SURABAYA	0	148587	3198	121	KAB. BANGKLAN	18	8127	781	66
KAB. SAMPANG	0	3768	166	4	KAB. LUMAJANG	0	11841	1068	20
KOTA KEDIRI	0	6326	401	0	KOTA PROBOLINGGO	0	9353	566	11
KAB. PASURUAN	70	12950	802	17	KAB. NGAWI	8	11280	990	18
KAB. PACITAN	16	10143	319	3	KOTA MALANG	0	32537	1317	86
KAB. MAGETAN	5	13779	1054	42	KOTA MADIUN	0	10866	578	2
KOTA BATU	0	5505	286	16	KAB. MOJOKERTO	0	12174	232	22
KAB. SITUBONDO	28	9079	902	4	KOTA BLITAR	1	8053	269	18
KAB. BLITAR	3581	14027	1641	40	KOTA PASURUAN	18	5218	258	7



## Lampiran 2

```
#load library
library(ggfortify) #Visualisasi scatter plot
library(corrplot) #Visualisasi Matriks Korelasi
library(GPArotation) #prasyarat faktor analysis
library(caret) #untuk cross validation
library(knitr)
install.packages("REdaS")
library(REdaS)
library(factoextra)
library(clValid)
library(tidyverse)
library(cluster)
library(psych)
library(Hmisc)

#import data
library(readr)
dataku <- read_delim("Data persebaran covid riska
1.csv",
                    delim =
";", escape_double = FALSE, trim_ws = TRUE)
View(dataku)
boxplot(dataku[,2:5])
summary(dataku)

# Pengecekan asumsi
bart_spher(dataku[,2:5])
KMO(dataku[,2:5])
cor(dataku[,2:5])
```

```
#menghitung vektor eigen dan nilai eigen
R <- cor(dataku[,2:5])
eigen<- eigen(R)
eigen$values #eigen dapat 2

#Melakukan PCA
pcadata<- prcomp(x = dataku[,2:5], scale. = TRUE,
center = TRUE)

#menentukan jumlah komponen utama
summary(pcadata)

#plot pca
plot(pcadata, type="lines")

#persamaan komponen utama
round(pcadata$rotation,2) #loading data
round(pcadata$sdev^2,2) #eigen
fviz_pca(pcadata) #Visualisasi data hasil rekonstruksi

#data pca fix
PCA_scoresbaru <- pcadata$x[,1:2]
View(PCA_scoresbaru)
new_Data=as.data.frame(PCA_scoresbaru)
new_Data
View(new_Data)
uji_bart(new_Data)
help(uji_bart)

library(fpc)
```

```

pamk.result <- pamk(PCA_scoresbaru)
pamk.result
pamk.result$nc
#menampilkan grafik sillhouette
fviz_nbclust(PCA_scoresbaru, pam, method =
"silhouette")
pam.hasil <- pam(PCA_scoresbaru, 2)
#jarak
pam.hasil$diss
#datakuframe hasil cluster
df.clusterbaruuuu =
data.frame(PCA_scoresbaru,pam.hasil$cluster)
View(df.clusterbaruuuu)
#Clustering
#sesuai abjad
table(pam.result$clustering, dataku$`Kab/Kota`)
#plot cluster
fviz_cluster(pam.hasil, data = PCA_scoresbaru)
# Validasi Cluster
intern <- clValid(PCA_scoresbaru, 2:5, clMethods =
c("hierarchical", "pam"),
validation = "internal")
summary(intern)
optimalScores(intern)
# Dendrogram
#single linkage
data.hccl <- PCA_scoresbaru %>%
  dist(method = "euclidean") %>%
  hclust(method = "single")
fviz_dend(data.hccl, k = 2,
          cex = 0.6, palette = "jco",

```

```

        rect = TRUE, rect_border = "jco", rect_fill =
TRUE,
        main = "Cluster Dendrogram Single Linkage")
#complete linkage
data.hcc2 <- PCA_scoresbaru %>%
  dist(method = "euclidean") %>%
  hclust(method = "complete")
fviz_dend(data.hcc2, k = 2,
          cex = 0.6, palette = "jco",
          rect = TRUE, rect_border = "jco", rect_fill =
TRUE,
          main = "Cluster Dendrogram Complete
Linkage")
#average
data.hcc3 <- PCA_scoresbaru %>%
  dist(method = "euclidean") %>%
  hclust(method = "average")
fviz_dend(data.hcc3, k = 2,
          cex = 0.6, palette = "jco",
          rect = TRUE, rect_border = "jco", rect_fill =
TRUE,
          main = "Cluster Dendrogram Average Linkage")
#ward
data.hcc4 <- PCA_scoresbaru %>%
  dist(method = "euclidean") %>%
  hclust(method = "ward.D2")
fviz_dend(data.hcc4, k = 2,
          cex = 0.6, palette = "jco",
          rect = TRUE, rect_border = "jco", rect_fill =
TRUE,
          main = "Cluster Dendrogram (Ward's Method)")
#pemilihan metode terbaik
m <- c("single","complete","ward", "average")

```

```
names(m) <- c("single","complete","ward", "average")
ac <- function(x){
  agnes(PCA_scoresbaru,method=x)$ac
}
map_dbl(m,ac)
# Deskripsi Statistik Cluster k-medoids
clust1 <- dataku[-c(2, 6, 11, 15, 27, 34 ),]
summary(clust1)
clust2 <- dataku[c(2, 6, 11, 15, 27, 34),]
summary(clust2)
# Deskripsi Statistik Cluster hierarchical
clust1_hieararki <- dataku[-c(22, 6, 11, 15, 27),]
summary(clust1)
clust2_hierarchical <- dataku[c(22, 6, 11, 15, 27),]
summary(clust2)
```

## Lampiran 3

## # Statistika Deskriptif

```
> summary(dataku)
```

Kab/Kota	Probable	Konfirmasi
Length:38	Min. : 0.00	Min. : 3348
Class :character	1st Qu.: 0.00	1st Qu.: 8850
Mode :character	Median : 0.00	Median : 11062
	Mean : 100.08	Mean : 17128
	3rd Qu.: 3.25	3rd Qu.: 17209
	Max. :3581.00	Max. :148587

Meninggal	Aktif
Min. : 166.0	Min. : 0.00
1st Qu.: 339.5	1st Qu.: 5.00
Median : 791.5	Median : 13.50
Mean : 860.2	Mean : 25.63
3rd Qu.:1064.5	3rd Qu.: 25.00
Max. :3198.0	Max. :157.00

## # Uji Bartlett's Test

```
> bart_spher(dataku[,2:5])
```

Bartlett's Test of Sphericity

```
Call: bart_spher(x = dataku[, 2:5])
```

```

X2 = 53.775
df = 6
p-value < 2.22e-16
```

## # Uji KMO

```
> KMO(dataku[,2:5])
```

Kaiser-Meyer-Olkin factor adequacy

```
Call: KMO(r = dataku[, 2:5])
```

```
Overall MSA = 0.62
```

```
MSA for each item =
```

Probable	Konfirmasi	Meninggal	Aktif
0.20	0.59	0.62	0.80

## # Matriks Multikolinieritas

```
> cor(dataku[,2:5])
```

	Probable	Konfirmasi	Meninggal	Aktif
Probable	1.00000000	-0.02391025	0.2135331	0.06741047
Konfirmasi	-0.02391025	1.00000000	0.7572529	0.64290471
Meninggal	0.21353310	0.75725292	1.00000000	0.55611065
Aktif	0.06741047	0.64290471	0.5561107	1.00000000

```
# Menentukan PCA
```

```
> summary(pcadata)
```

```
Importance of components:
```

	PC1	PC2	PC3	PC4
Standard deviation	1.524	1.0109	0.6764	0.44336
Proportion of Variance	0.581	0.2555	0.1144	0.04914
Cumulative Proportion	0.581	0.8365	0.9509	1.00000

```
# Menentukan Jumlah Cluster Optimum
```

```
> summary(intern)
```

```
Clustering Methods:
```

```
hierarchical pam
```

```
Cluster sizes:
```

```
2 3 4 5
```

```
Validation Measures:
```

		2	3	4	5
hierarchical Connectivity		2.9290	5.8579	11.2944	13.8778
Dunn		0.6691	1.1372	0.3129	0.3129
Silhouette		0.7884	0.7826	0.6238	0.5533

```
# Profilisasi Cluster Terbaik
```

```
# Cluster 1
```

```
> summary(clust1)
```

Kab/Kota	Probable	Konfirmasi
Length:32	Min. : 0.00	Min. : 3768
Class :character	1st Qu.: 0.00	1st Qu.: 9003
Mode :character	Median : 0.00	Median :11062
	Mean : 11.34	Mean :13681
	3rd Qu.: 4.25	3rd Qu.:15972
	Max. :3581.00	Max. :51106
Meninggal	Aktif	
Min. : 166.0	Min. : 0.00	
1st Qu.: 433.2	1st Qu.: 5.00	
Median : 803.5	Median : 14.00	
Mean : 830.1	Mean : 23.03	
3rd Qu.:1057.5	3rd Qu.: 23.00	
Max. :1886.0	Max. :157.00	

```
# Cluster 2
```

```
> summary(clust2)
```

Kab/Kota	Probable	Konfirmasi
Length:6	Min. : 0.000	Min. : 3348
Class :character	1st Qu.: 0.000	1st Qu.: 7506
Mode :character	Median : 0.000	Median : 10982
	Mean : 26.67	Mean : 35510
	3rd Qu.: 0.000	3rd Qu.: 27358
	Max. :16.000	Max. :148587

Meninggal		Aktif	
Min.	: 235.0	Min.	: 1.00
1st Qu.:	292.8	1st Qu.:	5.50
Median :	544.5	Median :	13.00
Mean	:1020.5	Mean	: 39.50
3rd Qu.:	1180.2	3rd Qu.:	67.75
Max.	:3198.0	Max.	:121.00



## RIWAYAT HIDUP



Rizka Allifatur Rohmah, lahir di Bojonegoro pada tanggal 21 Januari 2000. Memiliki nama panggilan Rizka. Bertempat tinggal di RT 01 RW 01 Desa Kanten, Kecamatan Trucuk, Kabupaten Bojonegoro. Merupakan anak Pertama dari 2 bersaudara dari Bapak Moh Roni dan Ibu Siti Mahfudhoh. Pendidikan yang pernah ditempuh yaitu MI Babul Ulum Kanten dan lulus pada tahun 2010. Menempuh Pendidikan di MTs Darussalam Kanten, lulus pada tahun 2013. Pada tahun yang sama melanjutkan pendidikan di MAN 2 Jombang dan lulus pada tahun 2016. Tahun 2016 melanjutkan studi di Universitas Islam Negeri Maulana Malik Ibrahim Malang menempuh Program Studi Matematika.



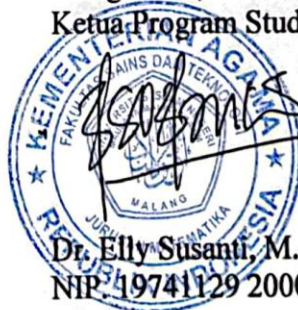
**KEMENTERIAN AGAMA RI**  
**UNIVERSITAS ISLAM NEGERI**  
**MAULANA MALIK IBRAHIM MALANG**  
**FAKULTAS SAINS DAN TEKNOLOGI**  
Jl. Gajayana No. 50 Dinoyo Malang Telp./Fax. (0341)558933

**BUKTI KONSULTASI SKRIPSI**

Nama : Rizka Allifatur Rohmah  
NIM : 16610110  
Fakultas / Program Studi : Sains dan Teknologi / Matematika  
Judul Skripsi : Penyebaran Covid-19 Pemprov Jawa Timur  
Menggunakan Analisis Cluster Hierarki Metode  
Ward dengan Pearson Correlation Distance  
Pembimbing I : Angga Dwi Mulyanto, M.Si  
Pembimbing II : Erna Herawati, M.Pd

No.	Tanggal	Hal	Tanda Tangan
1.	30 November 2022	Konsultasi Bab I, II dan III	1.
2.	05 Desember 2022	Konsultasi Kajian Keagamaan (Bab I dan II)	2.
3.	07 Maret 2023	Revisi Bab II	3.
4.	10 Maret 2023	ACC Kajian Keagamaan (Bab I-III)	4.
5.	05 Juni 2023	Konsultasi Bab IV dan V	5.
6.	26 Juni 2023	ACC Keseluruhan	6.
7.	26 Juni 2023	ACC Keseluruhan	7.

Mengetahui,  
Ketua Program Studi Matematika



Dr. Elly Susanti, M. Sc  
NIP. 19741129 200012 2 005