

## Aplikasi K-means cluster untuk pengelompokan unit kerja yang beresiko Covid-19: Studi kasus pada industri layanan kesehatan

Hasiholan Manurung<sup>1</sup>, Dias Irawati Sukma<sup>2</sup>

<sup>1,2</sup>Mahasiswa Program Studi Magister Teknik Industri, Universitas Mercu Buana, Jakarta

Corresponding author: [diasirawati2005@gmail.com](mailto:diasirawati2005@gmail.com)

---

**Abstrak.** Pandemi infeksi COVID-19 menjadi wabah yang sedang menjadi perhatian utama saat ini melebihi wabah demam berdarah yang saat ini masih terus terjadi. Fokus resiko terutama pada tenaga medis dan non-medis yang bekerja khususnya di pelayanan kesehatan menjadi sangat tinggi karena berhadapan langsung dengan pasien covid-19 atau pasien lain yang kemungkinan menderita namun belum terdeteksi dengan keluhan apenyakit lain. Terkait dengan ini karyawan yang bekerja di pelayanan kesehatan terpengaruh juga dengan kondisi kesehatan dan kebiasaan yang terus dilakukan sehari-hari berpengaruh terhadap tingkat resiko yang mungkin terjadi. Pasien yang positif covid-19 yang harus dirawat di rumah sakit biasanya memiliki tekanan respiratory yang akut yang bias berakibat fatal. Identifikasi pasien infeksi Covid-19 biasanya diperoleh setelah dilakukan pemeriksaan laboratorium yang dimana secara jumlah tidak sampai setengahnya didasari oleh penyakit bawaan (komorbid) seperti, diabetes, hipertensi dan yang terkait dengan jantung dan pembuluh darah.

Kata kunci: tenaga kesehatan, resiko, pelayanan kesehatan, cluster.

**Abstract.** *The pandemic of COVID-19 infection is an epidemic that is of major concern now more than the dengue epidemic that continues to occur today. The focus of risk, especially on medical and non-medical personnel working specifically in health services, is very high because it deals directly with co-19 patients or other patients who are likely to suffer but have not been detected with other illnesses. Related to this, the human resources working in the health service are affected as well as the health conditions and habits that are carried out daily that affect the level of risk that might occur. Covid-19 positive patients who must be hospitalized usually have acute respiratory distress which can be fatal. Identification of patients with Covid-19 infection is usually obtained after laboratory examinations which in less than half are based on congenital (comorbid) diseases such as diabetes, hypertension and those related to the heart and blood vessels.*

Keyword: health workers, risk, health service, cluster

---

### 1 Pendahuluan

Pemerintah Indonesia untuk pertama kali mempublikasikan kasus pandemi covid-19 pada Maret 2020 dengan kasus dugaan pada Warga Negara Asing yang masuk ke Indonesia yang berusia 31 dan 64 tahun. Namun pendapat lain mengatakan bahwa masuknya virus corona di Indonesia sudah sejak akhir Januari 2020, dengan adanya laporan ODP (Orang dalam Pantauan) yang diduga sebagai penyebab penyebaran virus corona secara lokal.

Menyebarnya virus corona mengharuskan tenaga medis untuk lebih waspada karena berhadapan langsung dengan penderita corona dan penderita penyakit lain yang mungkin menularkan namun belum terdeteksi. Resiko Covid-19 terhadap tenaga kesehatan, diangkat oleh penulis untuk meneliti lebih jauh terhadap tenaga kesehatan yang bekerja di Rumah Sakit XYZ. Dimana Rumah Sakit ini terdapat profesi yang beragam. Selain profesi Dokter dan Perawat terdapat juga 22 profesi tenaga kesehatan lain yang terlibat dalam jasa pelayanan Rumah Sakit.

Tujuan penelitian ini adalah menganalisis resiko Covid-19 terhadap 920 tenaga kesehatan dari 22 profesi tenaga kesehatan. Setelah dilakukan penyebaran kuisioner, di peroleh 759 responden yang memberikan inputan terhadap kuisioner yang disebar.

Tahapan clusterisasi meliputi pengelompokan asal unit kerja koresponden yang melibatkan profesi tenaga kesehatan dalam pelayanan. Berbagai masalah yang akan menjadi sumber resiko dikelompokkan lagi sesuai jabatan profesi tenaga kesehatan. Jabatan Profesi : Ahli Teknologi Laboratorium Medik, Audiologi, Biologi, Dietisien, Elektromedis, Fisikawan Medik, Fisioterapis, Okupasi Terapis, Penata Anestesi, Perekam Medis dan Informasi Kesehatan, Psikologi Klinis, Radiografer, Radioterapis, Refraksionis Optisien/ Optometris, Teknik Kardiovaskuler, Teknik Pelayanan Darah, Teknisi Gigi, Tenaga Sanitasi Lingkungan, Tenaga Teknis Kefarmasian, Terapis Gigi dan Mulut dan Terapis Wicara dengan jumlah total sebanyak 920 orang yang tersebar di 39 unit kerja di Rumah Sakit XYZ.

Penelitian tentang pengelompokan cluster unit kerja yang beresiko terinfeksi Covid-19 bagi tenaga kesehatan di Rumah Sakit, memang belum penulis temukan dalam jurnal lainnya, pengelompokan dengan K-Means Cluster pernah disampaikan oleh peneliti Rajabbi (2015) dengan membuat analisis resiko bagi developer pengembang perumahan di kota Mataram. Dari kelompok responden diketahui terbentuk 3 cluster dan masing masing cluster mengelompokkan faktor resiko yang dipilih responden. Disamping itu ada juga penelitian dari Edvira Rivanni (2009) tentang mengenai cara menganalisis dengan menggunakan cluste K-Means dengan tujuan melakukan pengelompokkan propinsi di Indonesia dalam produksi hasil pertanian (padi, jagung, dan kedelai).

Adapun kuisoner diberikan kepada berbagai profesi di RS. XYZ melihat profesi ini melakukan pekerjaan keprofesiannya dengan berhadapan langsung ataupun tidak namun masih dalam lingkup lingkungan perawatan kesehatan memiliki resiko yang lebih besar dibandingkan dengan profesi di luar lingkup yang bukan dibidang pelayanan kesehatan. Sebab itu didalam melakukan kegiatan keprofesiannya sehari-hari harus memiliki sikap kehati-hatian yang lebih disbanding dengan profesi lainnya. Karena itu disini kami tertarik untuk mengangkat topik ini karena tenaga medis dan pendukung lainnya yang bekerja di pelayanan kesehatan khususnya di RS. XYZ menjadi pilar utama didalam menghadapi pandemi Covid-19 sehingga bias menjadi semacam acuan untuk profesi lain untuk mencegah resiko terjangkau virus dengan melakukan kebiasaan yang dianggap perlu dan menghindari kebiasaan hidup yang tidak perlu.

## 2 Kajian Teori

### **Coronavirus-19**

Patogen yang menjadi penyebab penyakit corona yang adalah coronavirus (CoV) sangat berbahaya karena sangat cepat masuk dan menyerang secara kelompok besar di dalam masyarakat karena masuk dan menyerang melalui sistem pernapasan manusia termasuk golongan (SARS)-CoV dan (MERS)-CoV. Dimulai pada Desember 2019, sekelompok pasien dirawat di rumah sakit dengan diagnosis awal pneumonia dari etiologi tidak diketahui penyebabnya. Pasien yang terinfeksi ini secara epidemiologis terkait dengan penjualan grosir dari pasar makanan laut dan hewan liar basah di Wuhan, Provinsi Hubei, Cina. Diinformasikan pada laporan awal perkiraan permulaan wabah Coronavirus yang sangat potensial penularan infeksi. Diperkirakan, sebagai penyebabnya adalah bertambahnya jumlah reproduksi untuk Novel 2019 Coronavirus (COVID-19) yang diberi nama oleh WHO pada 11 Februari 2020.

Infeksi yang dialami pasien yang dirawat di rumah sakit dengan sindrom tekanan re-spiratory akut dan penyebab orang yang meninggal terinfeksi COVID-19. Pasien yang dirawat inap di rumah sakit diidentifikasi telah memiliki infeksi COVID-19 yang telah dikonfirmasi laporan pemeriksaannya dari laboratorium, kurang dari setengah pasien memiliki penyakit yang mendasari penyakit bawaan (komorbid), termasuk diabetes, hipertensi, dan penyakit kardiovaskular. Orang yang terinfeksi diduga terinfeksi dari rumah sakit, kemungkinan lainnya karena tertular atau infeksi nosokomial. COVID-19 bukan merupakan virus dengan penyebaran super panas yang ditularkan dari pasien ke orang, tetapi cenderung penularannya dari orang yang terinfeksi di berbagai lokasi di seluruh rumah sakit melalui mekanisme yang tidak diketahui. Diketahui hanya pasien yang telah terkonfirmasi uji klinis yang sakit, sehingga ada kemungkinan lebih banyak pasien atau orang yang mungkin terinfeksi.

Periode infeksi COVID-19 yang muncul tergantung pada usia pasien dan status sistem kekebalan tubuh pasien. Masa inkubasi sekitar 5,2 hari bagi orang yang memiliki gejala terinfeksi COVID-19. Gejala COVID-19 hingga penyebab kematian berkisar antara 6 hingga 41 hari dengan median 14 hari rentang usia pasien > 50 tahun lebih rentan dibandingkan dengan mereka yang berusia di bawah 50. Gejala paling umum pada permulaan periode awal infeksi penyakit COVID-19 adalah demam, batuk, dan kelelahan, sementara gejala lainnya termasuk produksi dahak, sakit kepala, hemoptisis, diare, dyspnoea, dan limfopenia. oleh CT scan dada disajikan sebagai pneumonia. Penularan dari orang ke orang terjadi terutama melalui kontak langsung atau melalui tetesan yang disebarkan oleh batuk atau bersin dari orang yang terinfeksi.

**Clustering** atau pengelompokan merupakan proses menggabungkan obyek kedalam kelompok obyek yang hampir sama. Proses pengclusteran merupakan satu dari sekian banyak fungsi proses data mining dengan tujuan menemukan kelompok yang akan diidentifikasi dengan obyek yang hampir sama. Hal penting yang perlu dilakukan dalam proses pengelompokan (pengclusteran) dengan cara menyatakan sekumpulan pola ke dalam kelompok yang sesuai guna menemukan kesamaan dan perbedaan sehingga dapat menghasilkan kesimpulan yang berharga.

**Algoritma K-means**

Algoritma K-means merupakan salah satu metode data dari klustering non hirarki yang berusaha mempartisi data ke dalam satu atau lebih cluster atau kelompok. Penggunaan metode K-Means dengan mempartisi data ke dalam cluster / kelompok sehingga data karakteristik yang sama (*High intra class similarity*) dikelompokkan ke dalam satu cluster yang sama dan yang memiliki karakteristik yang berbeda (*Low inter class similarity*) dikelompokkan pada kelompok yang lain. Dalam melakukan proses klustering pertama kali dimulai dengan mengidentifikasi data yang akan dikluster,  $X_{ij}$  ( $i=1, \dots, n$ ;  $j=1, \dots, m$ ) dengan  $n$  adalah jumlah data yang akan dikluster dan  $m$  adalah jumlah variabel. Pada awal iterasi, pusat setiap cluster ditetapkan secara bebas (sembarang),  $C_{kj}$  ( $k=1, \dots, k$ ;  $j=1, \dots, m$ ). Proses selanjutnya menghitung jarak antara setiap data dengan setiap pusat cluster. Untuk melakukan penghitungan jarak data ke- $i$  ( $x_i$ ) pada pusat cluster ke- $k$  ( $c_k$ ), diberi nama (**dik**), dapat digunakan formula Euclidean [2] seperti pada persamaan (1), yaitu:

$$d^{ik} = \sqrt{\sum_{j=1}^m (C_{ij} - C_{kj})^2} \dots\dots\dots(1)$$

Sekelompok data akan menjadi anggota dari cluster ke- $k$  apabila jarak data tersebut ke pusat cluster ke- $k$  bernilai paling kecil apabila dibandingkan dengan jarak ke pusat cluster lainnya. Dapat dihitung dengan menggunakan persamaan (2) Selanjutnya, kelompokkan data-data yang menjadi anggota pada setiap cluster.

$$\text{Min } \sum_{k=1}^k d_{ik} = \sqrt{\sum_{j=1}^m (C_{ij} - C_{kj})^2} \dots\dots\dots(2)$$

Pusat cluster yang baru memiliki nilai yang dapat dihitung dengan cara mencari nilai rata-rata dari data-data yang menjadi anggota pada cluster tersebut, dengan menggunakan rumus pada persamaan 3:

$$C_{kj} = \frac{\sum_{t=1}^p X_{tj}}{p} \dots\dots\dots(3)$$

Dimana:  
 $x_{ij} \in$  kluster ke -  $k$   
 $p$  = banyaknya anggota cluster ke  $k$

Proses algoritma dasar dalam k-means meliputi:

1. Penentuan jumlah cluster ( $k$ ), tetapkan pusat cluster sembarang.
2. Menghitung jarak setiap data ke pusat cluster menggunakan persamaan (1).
3. Mengelompokkan data ke dalam cluster yang dengan jarak yang paling pendek menggunakan persamaan (2).
4. Menghitung pusat cluster yang baru menggunakan persamaan (3)
5. Terakhir, ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke cluster yang lain.

### 3 Metode

Penelitian ini tergolong penelitian deskriptif kuantitatif. Data yang dikumpulkan terdiri dari data primer dan data sekunder. Data sekunder diperoleh melalui penyebaran kuesioner terhadap 22 profesi tenaga kesehatan yang bekerja pada industri layanan kesehatan pada kesus RS XYZ pada bulan Juni tahun 2020. Jumlah responden yang digunakan pada penelitian ini sebanyak 759 responden, telah memenuhi syarat ideal analisis faktor (Gunartha, 2013) minimal 50 responden. Responden seimbang antara keterwakilan responden yang bekerja pada lingkup kegiatan lapangan dan kegiatan internal (kantor) developer. Pengelolaan data analisis faktor, analisis cluster, dan perhitungan kecenderungan sentral data menggunakan alat bantu komputer dengan software SPSS (*Statistical Product and Service Solution*).

Adapun kuisoner diberikan kepada berbagai profesi di RS. XYZ melihat profesi ini melakukan pekerjaan keprofesiannya dengan berhadapan langsung ataupun tidak namun masih dalam lingkup lingkungan perawatan kesehatan memiliki resiko yang lebih besar dibandingkan dengan profesi di luar lingkup yang bukan dibidang pelayanan kesehatan. Sebab itu di dalam melakukan kegiatan keprofesiannya sehari-hari harus memiliki sikap kehati-hatian yang lebih dibanding dengan profesi lainnya. Karena itu disini kami tertarik untuk mengangkat topik ini karena tenaga medis dan pendukung lainnya yang bekerja di pelayanan kesehatan khususnya di RS. XYZ menjadi pilar utama didalam menghadapi pandemi Covid-19 sehingga bias menjadi semacam acuan untuk profesi lain untuk mencegah resiko terjangkit virus dengan melakukan kebiasaan yang dianggap perlu dan menghindari kebiasaan hidup yang tidak perlu.

Tahapan penelitian dapat diuraikan sebagai berikut:

#### 1. Penetapan variabel awal

Pengidentifikasian variabel resiko berdasarkan studi literatur yang berkaitan, kemudian digunakan pada kuesioner pendahuluan. Kuesioner pendahuluan bertujuan untuk menetapkan variabel awal resiko yang memiliki nilai yang relevan dan mendapatkan tambahan variabel resiko yang sesuai dengan kondisi sesungguhnya dari para responden yang berpengalaman. Variabel awal resiko digunakan untuk kuesioner utama jika memenuhi syarat jawaban kumulatif responden "Ya"  $\geq 1$  (satu) maka digunakan, dan apabila jawaban kumulatif "Tidak" = 0 (nol), maka variabel tidak digunakan.

#### 2. Uji validitas dan uji reliabilitas

3. Ditolak bila  $r_{hitung} > r_{\alpha}$ , yang menyimpulkan bahwa variabel tersebut valid (Sanusi, 2005).

4. Uji reliabilitas (keterpercayaan) menunjuk pada pengertian apakah sebuah instrument dapat mengukur sesuatu yang diukur secara konsisten dari waktu ke waktu. Dapat menggunakan koefisien Cronbach Alpha ( $\alpha$ ), instrumen memiliki tingkat reliabilitas yang tinggi jika memberikan nilai Cronbach Alpha ( $\alpha$ )  $> 0,70$  (Ghozali, 2001).

#### Tahapan Analisis Faktor

Tahapan-tahapan dalam analisis faktor adalah sebagai berikut:

- Memilih variabel yang diteliti. Jumlah variabel yang dipilih sangat bergantung pada permasalahannya dan dituangkan dalam kuesioner.
- Bentuk matriks korelasi. Matriks data mentah yang berisi kuesioner diubah menjadi matriks korelasi. Dalam matriks korelasi variabel-variabel yang diukur mempunyai unit dan skala pengukuran yang berbeda. Penggunaan matriks ini untuk menghilangkan perbedaan yang diakibatkan oleh mean dan diversifikasi variabel.
- Uji asumsi. Untuk menguji ketepatan dalam analisis faktor, uji statistik yang digunakan adalah Barlett's Test Sphericity, Kaiser-Meyer-Olkin (KMO) dan Measure of Sampling Adequacy (MSA). Uji Bartlett's merupakan tes statistik untuk menguji apakah variabel-variabel bebas yang dilibatkan berkorelasi.  $H_0$  = Tidak ada korelasi antar variabel bebas,  $H_1$  = Ada korelasi antar variabel bebas. Kriteria uji dengan melihat p-value (signifikansi) yaitu terima  $H_0$  jika Sig.  $> 0.05$  atau tolak  $H_0$  jika Sig.  $< 0.05$ . Nilai KMO merupakan indeks untuk membandingkan besarnya koefisien korelasi dengan besarnya koefisien korelasi parsial. Kriteria penilaian uji KMO dari matriks antar variabel adalah bila nilai KMO  $> 0.5$  artinya data cukup untuk analisis faktor dan nilai KMO. MSA (*Measures of Sampling Adequacy*) untuk mengukur kecukupan sampling untuk setiap variabel individual. Analisis faktor hanya dapat dilakukan jika nilai MSA  $\geq 0,5$  artinya variabel masih bisa diprediksi dan bisa diprediksi lebih lanjut dan MSA  $\leq 0.5$  artinya variabel tidak dapat diprediksi dan tidak dapat dianalisis lebih lanjut atau harus dikeluarkan (Santoso, 2002).

- d. Proses ekstraksi dan rotasi faktor. Pada penelitian ini analisis faktor didasarkan pada matriks kovarian karena seluruh variabel memiliki satuan yang sama. Hal yang dilakukan sebelum mengekstraksi faktor yaitu membentuk matriks kovarian. Rotasi faktor tujuannya untuk menyederhanakan struktur faktor agar lebih mudah dalam menginterpretasikannya. Ada beberapa metode rotasi, namun karena tujuan penelitian ini untuk mengurangi jumlah variabel awal, maka pilihan rotasi yang cocok adalah rotasi orthogonal, yang dibedakan quartimax, varimax dan equamax.
- e. Interpretasi faktor. Pemberian nama baru pada faktor-faktor yang terbentuk yang dianggap bisa mewakili variabel-variabel anggota faktor tersebut.
- f. Menghitung factor scores. Melakukan perhitungan factor scores untuk masing-masing responden. *Factor scores* akan digunakan sebagai data input pada analisis cluster.

#### Penilaian Resiko

Rumus sederhana untuk penilaian resiko (R) (Zhi, 1995) adalah:  $R = \text{Frekuensi (F)} \times \text{Dampak (I)}$ . Kemudian penentuan nilai frekuensi dan dampak resiko menggunakan perhitungan nilai rata-rata (Average Index) yang dirumuskan oleh Majid dan McCaffer (1997) sebagai berikut:

$$\text{Average Index (AI)} = \frac{\sum_{i=1}^5 a_i \cdot x_i}{N} \dots\dots\dots(4)$$

Dengan  $a_i$  = konstanta penilai (1,2,3,4,5)

$x_i$  = frekuensi responden

$N$  = total jumlah responden

Kriteria nilai frekuensi dan dampak resiko berdasarkan nilai skala resiko rendah, sedang dan tinggi terhadap variabel responden sebagai berikut; usia, pernah keluar rumah/tempat umum pernah menggunakan transportasi umum, pernah melakukan perjalanan keluar kota/ internasional, pernah mengikuti kegiatan yang melibatkan orang banyak, memiliki riwayat kontak dengan orang yang dinyatakan ODP,PDP atau Konfirm Covid-19, pernah mengalami demam/batuk/pilek/sakit tenggorokan/sesak dalam 14 hari terakhir, memiliki Penyakit Komorbid (penyerta) dengan menggunakan kuesioner, di analisis dampak resiko kuesioner pertanyaan kepada populasi tenaga kesehatan di Rumah Sakit XYZ.

#### Tahapan Analisa Cluster

Input data dalam analisis cluster merupakan nilai skor faktor. Secara umum tahapan dalam analisis cluster yaitu:

- a. Penetapan tujuan cluster.
- b. Pembentukan cluster dengan metode K-Means
- c. Interpretasi dan memprofil cluster.

## 4 Hasil dan Pembahasan

#### Pembentukan Cluster Pertama pada K-means Cluster

Dengan menggunakan bantuan *software* SPSS 22, diperoleh output seperti pada Tabel 1. Tabel *initial cluster* tersebut menunjukkan tiga buah cluster yang pertama kali terbentuk. Selanjutnya dengan metode K-means cluster dilakukan pengujian untuk melakukan realokasi cluster yang ada. Proses tersebut dinamakan *Iteration*, yang memuat perubahan pada *initial cluster* atau disebut *iteration history* yang dapat dilihat pada Tabel 2.

**Tabel 1 Initial Cluster Centers**

	Cluster		
	1	2	3
Zscore: Usia	-,38253	2,61070	-,38253
Zscore: Tidak.Keluar_Rumah	-,40923	-,40923	2,44038
Zscore: Transportasi	-,84403	1,18324	1,18324
Zscore: Perjalanan	-3,35069	,29805	-3,35069
Zscore: Kegiatan	-1,89612	,52670	,52670
Zscore: RiwayatKontak	-1,67641	,59573	,59573
Zscore: Demam.Batuk.Sakittenggorokan	-2,77740	,35957	,35957
Zscore: Komorbid	2,32917	2,32917	-,42877
Zscore: Jenis Penyakit	-2,07413	-2,07413	,48149

Hasil akhir proses iterasi dari K-means atau Final Cluster Centers menghasilkan tiga cluster untuk sembilan aktivitas yang dilakukan tenaga kesehatan di unit kerja di RS XYZ tahun 2020 selama pandemi Covid-19.

**Tabel 2 Iteration History<sup>a</sup>**

Iteration	Change in Cluster Centers		
	1	2	3
1	3,643	2,582	4,320
2	,391	,126	,057
3	,623	,055	,114
4	,905	,221	,108
5	,162	,078	,008
6	,055	,025	,000
7	,000	,000	,000

**Analisis Pengelompokan Aktivitas Tenaga Kesehatan RS. XYZ**

Dari empat variabel yang relevan, untuk membedakan isi cluster, dapat dianalisis aktivitas yang termasuk cluster 1, cluster 2, dan cluster 3. Pengelompokan 9 aktivitas disajikan dalam Tabel 3.

**Tabel 3 Hasil pengelompokan 9 variabel**

Cluser	Jumlah	Persentase	Unit Kerja
1	4002	27.50 %	Departemen Medik Gigi dan Mulut, Instalasi Pengelolaan Logistik, Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Instalasi Farmasi, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Unit Fasilitas Medik, Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV, Unit Pelayanan Terpadu Kencana, Unit Pengadaan Barang/Jasa, Unit Produksi Makanan, Unit Rawat Jalan Terpadu, Unit Sanitasi dan Lingkungan, Departemen Medik Kebidanan dan Penyakit Kandungan.
2	7284	50.13 %	Departemen Medik Gigi dan Mulut, Instalasi Pengelolaan Logistik, Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Instalasi Farmasi, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Unit Fasilitas Medik, Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV, Unit Pelayanan Terpadu Kencana, Unit Pengadaan Barang/Jasa, Unit Produksi Makanan, Unit Rawat Jalan Terpadu, Unit Sanitasi dan Lingkungan, Departemen Medik Kebidanan dan Penyakit Kandungan.
3	3243	22.32 %	Departemen Medik Gigi dan Mulut, Instalasi Pengelolaan Logistik, Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Instalasi Farmasi, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Unit Fasilitas Medik, Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV, Unit Pelayanan Terpadu Kencana, Unit Pengadaan Barang/Jasa, Unit Produksi Makanan, Unit Rawat Jalan Terpadu dan Unit Sanitasi dan Lingkungan.

Berdasarkan Tabel 3 di atas terlihat bahwa pengelompokan 9 aktivitas ke dalam tiga cluster adalah sebagai berikut:

- Cluster 1 dengan prosentase 27,50 % terdiri dari usia saat ini, Apakah anda pernah keluar rumah/ tempat umum (pasar, fasinkes, kerumunan orang, dan lain-lain ? Apakah pernah menggunakan transportasi umum?
- Cluster 2 dengan prosentase 50,13% terdiri Apakah pernah melakukan perjalanan keluar kota/ internasional? (Wilayah yng terjangkau/ zona merah), Apakah anda mengikuti kegiatan yang melibatkan orang banyak? Apakah memiliki riwayat kontak erat denangan orang yang dinyatakan ODP,PDP atau Konfirm Covid-19 (berjabat tangan, berbicara, berada dalam satu ruangan/ satu rumah)?,
- Cluster 3 dengan prosentase 22,32% terdiri dari Apakah pernah mengalami demam/ batuk/ pilek/ sakit tenggorokan/ sesak dalam 14 hari terakhir ?, Penyakit Komorbid (penyerta) yang diderita, Jenis Penyakit Komorbid (penyerta).

#### Interpretasi Karakteristik Klasifikasi Aktifitas

Interpretasi terhadap ketiga cluster yang terbentuk dimulai dengan menganalisis variabel-variabel yang membedakan antara tiga cluster. Analisis ini dilakukan dengan melihat nilai signifikansi pada tabel ANOVA. Apabila nilai signifikansi > 0.05, maka tidak ada perbedaan yang berarti antara cluster 1, cluster 2 dan cluster 3 dan jika sebaliknya, nilai signifikansi < 0.05 maka ada perbedaan yang berarti antara cluster 1, cluster 2 dan cluster 3, yang berhubungan dengan variabel tersebut. Tabel 4 berikut menyajikan variabel-variabel pembeda cluster.

**Tabel 4** ANOVA uji signifikansi variabel pembeda Cluster

Cluster	Mean Square	df	Error	Mean Square	df	F	Sig.
	159.726	2	.580	756		275.346	.000
	.726	2	1.001	756		.725	.485
	2.560	2	.996	756		2.571	.077
	3.471	2	.993	756		3.494	.031
	1.125	2	1.000	756		1.125	.325
	6.363	2	.986	756		6.455	.002
	374.106	2	.013	756		28894.161	.000
	146.853	2	.614	756		239.118	.000
	192.811	2	.493	756		391.444	.000

Berdasarkan tabel anova di atas, terlihat bahwa angka signifikan untuk semua variabel aktifitas usia, keluar rumah, pakai transportasi umum, keluar kota, kegiatan dengan orang banyak, kontak dengan ODP, PDP, Confirm Covid, sakit demam, ISPA dalam 14 hari, komorbid diderita, komorbid lebih kecil dari 0.05. Hal ini menunjukkan bahwa ada perbedaan yang berarti antara cluster 1, cluster 2, dan cluster 3, yang berhubungan dengan semua variabel tersebut. Semakin besar angka F maka semakin besar perbedaan cluster 1, cluster 2, dan cluster 3 untuk variabel yang bersangkutan. Berdasarkan Tabel 3 di atas, terlihat angka F terbesar adalah 275, 346. Hal ini menunjukkan tenaga kesehatan dalam cluster 1 faktor nilai resiko yang sangat berbeda bila dibandingkan dengan faktor resiko cluster 2, cluster 1 berakibat tingkat resiko yang sangat berbeda bila dibandingkan dengan jumlah produksi padi cluster 3, dan cluster 2 memiliki tingkat faktor resiko yang sangat berbeda bila dibandingkan dengan tingkat resiko pada cluster 3. Setelah diketahui variabel-variabel yang relevan untuk membedakan cluster 1, cluster 2, dan cluster 3, selanjutnya dianalisis tingkat nilai resiko masing-masing tenaga kesehatan yang termasuk ke dalam masing-masing cluster yang dapat dilihat dari nilai akhir pusat cluster pada variabel yang membedakan.

**Tabel 5** Hasil akhir karakteristik masing-masing Cluster

Variabel	Cluster		
	1	2	3
Usia	-.31292	1.13000	-.38253
Keluar Rumah	-.11102	.04549	.00205
Pakai Transportasi umum	.05174	-.14311	.04630
Keluar Kota	-.25350	.08456	.01217
Kegiatan dgn orang banyak	-.14944	.03698	.01216
Kontak dengan ODP,PDP, confirm Covid	-.35540	.08812	.02886
Sakit demam, ISPA dlm 14 hari	-2.77740	.34289	.35957
Komorbid diderita	.14847	1.03822	-.42877
Komorbid	-.08312	-1.20413	.48149

Berdasarkan Tabel 5 final *cluster center* di atas dapat disimpulkan bahwa dari sembilan variabel yang relevan untuk membedakan isi cluster dalam mengelompokkan resiko yang termasuk cluster 1, cluster 2, dan Cluster 3.

a. Variabel *Resiko Usia*

Rata-rata tingkat resiko untuk cluster 3 (-.38253) < cluster 1(.31292) < cluster 2 (1.13000). Berarti pada tahun 2020 tingkat resiko cluster 2 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 3 memiliki tingkat resiko jumlah yang paling sedikit.

b. Variabel *Keluar Rumah*

Rata-rata tingkat resiko untuk cluster 1 (-.11102) < cluster 3 (.00205) < cluster 2 (.04549). Berarti berarti pada tahun 2020 tenaga kesehatan pada cluster 3 memiliki resiko paling tinggi dibandingkan dengan tingkat resiko pada cluster lainnya, dimana cluster 1 memiliki resiko dengan tingkat yang paling rendah.

c. Variabel *Resiko Pakai Transportasi Umum*

Rata-rata tingkat resiko untuk cluster 2 (-.14311) < cluster 3(.04630) < cluster 1 (.05174). Berarti pada tahun 2020 tingkat resiko cluster 1 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 2 memiliki tingkat resiko jumlah yang paling sedikit.

d. Variabel *Resiko Keluar Kota*

Rata-rata tingkat resiko untuk cluster 1 (-.25350) < cluster 3(.01217) < cluster 2 (.08456). Berarti pada tahun 2020 tingkat resiko cluster 2 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 1 memiliki tingkat resiko jumlah yang paling sedikit.

e. Variabel *Resiko Kegiatan Dengan Orang Banyak*

Rata-rata tingkat resiko untuk cluster 1 (-.14944) < cluster 3(.01216) < cluster 2 (.03698). Berarti pada tahun 2020 tingkat resiko cluster 2 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 1 memiliki tingkat resiko jumlah yang paling sedikit.

f. Variabel *Resiko Kontak dengan ODP, PDP, Confirm Covid*

Rata-rata tingkat resiko untuk cluster 1 (-.35540) < cluster 3(.02886) < cluster 2 (.08812). Berarti pada tahun 2020 tingkat resiko cluster 2 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 1 memiliki tingkat resiko jumlah yang paling sedikit.

g. Variabel *Resiko Sakit Demam, ISPA dalam 14 hari*

Rata-rata tingkat resiko untuk cluster 1 (-2.77740) < cluster 2(.34289) < cluster 3 (.35957). Berarti pada tahun 2020 tingkat resiko cluster 3 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 1 memiliki tingkat resiko jumlah yang paling sedikit.

h. Variabel *Resiko Komorbid diderita*

Rata-rata tingkat resiko untuk cluster 3 (-.42887) < cluster 1(.14847) < cluster 2 (1.03822). Berarti pada tahun 2020 tingkat resiko cluster 2 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 3 memiliki tingkat resiko jumlah yang paling sedikit.

- i. Variabel Resiko Komorbid  
Rata-rata tingkat resiko untuk cluster 2 (-1.20413) < cluster 1(.08312) < cluster 3(.48149). Berarti pada tahun 2020 tingkat resiko cluster 3 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 2 memiliki tingkat resiko jumlah yang paling sedikit.

## 5 Kesimpulan

1. Rata-rata tingkat resiko usia cluster 1 < cluster 2 < cluster 3. Hal ini berarti pada tahun 2020 tingkat resiko usia cluster 3 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 1 memiliki tingkat resiko yang paling sedikit.
2. Rata-rata tingkat resiko pakai transportasi umum untuk cluster 2 < cluster 3 < cluster 1. Hal ini berarti pada tahun 2020 tingkat resiko cluster 1 paling tinggi dibandingkan dengan tingkat resiko cluster lainnya, dimana pada cluster 2 memiliki tingkat resiko jumlah yang paling sedikit.
3. Cluster yang terbentuk dari pengelompokan variabel ada tiga, yaitu:
  - Cluster 1 sebanyak 27.50 % tenaga kesehatan yang bekerja di area Departemen Medik Gigi dan Mulut, Instalasi Pengelolaan Logistik Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Instalasi Farmasi, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Unit Fasilitas Medik, Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV, Unit Pelayanan Terpadu Kencana, Unit Pengadaan Barang/Jasa, Unit Produksi Makanan, Unit Rawat Jalan Terpadu, Unit Sanitasi dan Lingkungan, Departemen Medik Kebidanan dan Penyakit Kandungan.
  - Cluster 2 sebanyak 50.13 % tenaga kesehatan yang bekerja di area Departemen Medik Gigi dan Mulut, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Instalasi Farmasi, Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV. Unit Produksi Makanan, Unit Rawat Jalan Terpadu, Unit Sanitasi dan Lingkungan, Departemen Medik Kebidanan dan Penyakit Kandungan.
  - Cluster 3 sebanyak 22.32 % tenaga kesehatan yang bekerja di area Departemen Medik Gigi dan Mulut, Instalasi Pengelolaan Logistik Instalasi Pelatihan dan Simulasi Klinik (ICTEC), Instalasi Farmasi, Departemen Medik Kesehatan Jiwa (Psikiatri), Departemen Medik Patologi Anatomi, Departemen Medik Radioterapi, Departemen Medik Rehabilitasi Medik, Unit Fasilitas Medik, Unit Pelayanan Onkologi Radiasi, Unit Pelayanan Terpadu HIV, Unit Pelayanan Terpadu Kencana, Unit Pengadaan Barang/Jasa, Unit Produksi Makanan, Unit Rawat Jalan Terpadu dan Unit Sanitasi dan Lingkungan.

## Referensi

- Hussin.A Rothan, Siddapa.N, Byarareddi, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, *Journal of Autoimmunity* (2020)
- Bogoch, A. Watts, A. Thomas-Bachli, C. Huber, M.U.G. Kraemer, K. Khan, Pneumonia of unknown etiology in Wuhan, China: potential for international spread via commercial air travel, *J. Trav. Med.* (2020), <https://doi.org/10.1093/jtm/taaa008>.
- H. Lu, C.W. Stratton, Y.W. Tang, Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle, *J. Med.* (2020) 401–402, <https://doi.org/10.1002/jmv.25678>.
- Rajabi Mubarak, Suryawan Murtiadi, Heri Sulistiyono, pengembangan kriteria analisis resiko bagi developer perumahan di Kota Mataram, (2015),
- Edmira Rivani (2009). Aplikasi K-Means Cluster Untuk pengelompokan provinsi berdasarkan produksi padi, jagung, kedelai dan kacang hijau tahun 2009. Jurusan Statistika Terapan Universitas Padjajaran Bandung.
- Jananto, Arief. 2010. *Memprediksi Kinerja Mahasiswa Menggunakan Teknik Data Mining (Studi kasus data akademik mahasiswa UNISBANK)*. Tesis Tidak Terpublikasi. Yogyakarta: Universitas Gajah Mada.
- Han, Jiawei; & Kamber, Micheline. 2001. *Data Mining Concepts and Techniques Second Edition*. San Francisco: Morgan Kaufman.
- Giyanto, Heribertus. 2008. *Penerapan algoritma Clustering K-Means, K-Medoid, Gath Geva*. Tesis Tidak Terpublikasi. Yogyakarta: Universitas Gajah Mada.
- Hair, J.F. Jr., Anderson, R.E., Tatham, R.L., dan Black W. C. (1992). *Multivariate Data Analysis, Fifth Edition*. New Jersey: Prentice Hall.