



LEMBAR JAWABAN FINAL 1



SEBELAS MARET STATISTICS OLYMPIAD 2020

NOMOR PESERTA

SSF118308

NAMA TIM

arah



@ihf8891p



@ssfuns



ssf.uns.ac.id



BAB I

RANCANGAN PENYELESAIAN MASALAH

1. *Problem Statement*

Masalah :

Mempercepat penanganan COVID-19 yang lebih terarah untuk semua provinsi di Indonesia

Tujuan :

1. Menentukan kelompok/cluster provinsi di Indonesia berdasarkan data
2. Menentukan saran kinerja pemerintah dalam penanganan COVID-19 yang lebih efektif dan efisien berdasarkan karakteristik dari setiap klaster

2. *Dasar Teori Clustering*

Dalam mencapai tujuan masalah tersebut, digunakan metode *Clustering* dalam *Machine Learning*. *Clustering* adalah pengelompokan data dengan sifat yang mirip. Data untuk *clustering* tidak memiliki label (kelas). Dalam penelitian ini digunakan 2 jenis clustering yaitu *K-Means Clustering* dan *Agglomerative Hierarchical Clustering*

a. K-Means Clustering

Algoritma K-means mengelompokkan data menjadi sebanyak K kelompok sesuai yang kita definisikan. Algoritma ini disebut juga sebagai *flat clustering*, artinya kelompok satu memiliki kedudukan sejajar dengan k . Algoritma dari K-Means Clustering adalah sebagai berikut :

1. Tentukan sebanyak K kelompok yang kita inginkan.
2. Inisiasi centroid untuk setiap kelompok.
3. Hitung kedekatan suatu data terhadap *centroid*, kemudian masukkan data tersebut ke kelompok yang centroid-nya memiliki sifat terdekat dengan dirinya.
4. Pilih kembali centroid untuk masing-masing kelompok, yaitu dari anggota kelompok tersebut.
5. Ulangi langkah-langkah sebelumnya sampai tidak ada perubahan anggota untuk semua kelompok

b. Agglomerative Hierarchical Clustering

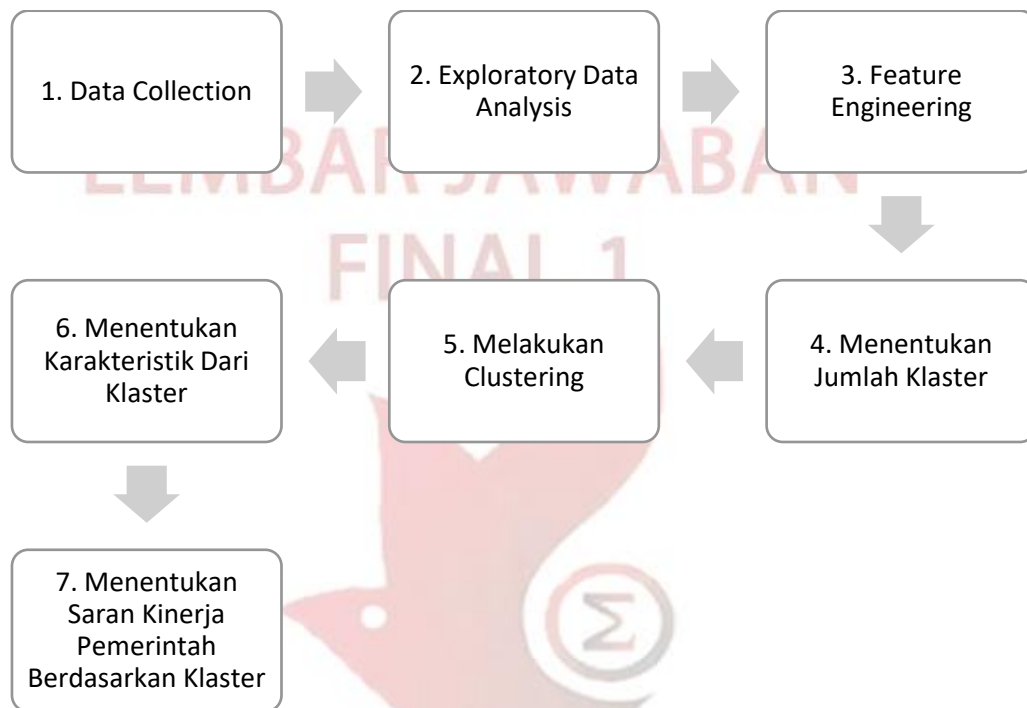
Hierarchical clustering adalah teknik untuk membentuk pembagian bersarang (*nested partition*). Berbeda dengan K-means yang hasil *clustering*-nya berbentuk *flat* atau rata, *hierarchical clustering* memiliki satu *cluster* paling atas yang mencakup konsep seluruh *cluster* dibawahnya. Agglomerative Hierarchical



Clustering dimulai dari beberapa *flat clusters*; pada setiap langkah iterasi, kita menggabungkan dua *clusters* termirip.

3. Workflow Penyelesaian Masalah

Sehingga, secara garis besar, diagram alir pemodelan clustering ini adalah sebagai berikut :



1. Data Collection : Mengumpulkan data dan informasinya.
2. Exploratory Data Analysis : Melakukan eksplorasi data untuk memahami dan mendapatkan gambaran dari data yang dimiliki untuk pemodelan berikutnya dengan statistika deskriptif dan visualisasi data.
3. Feature Engineering : Mengolah feature dengan memilih feature yang akan digunakan dalam pemodelan data dan mengisi data yang kosong (*missing value*)
4. Menentukan Jumlah Klaster
 - a. Metode *Elbow*

Yaitu menentukan jumlah *cluster* terbaik dengan cara melihat perbandingan antara jumlah *cluster* (K) dengan SSE. Jika nilai *cluster* pertama dengan nilai *cluster* kedua memberikan sudut siku dalam grafik atau nilainya mengalami penurunan paling besar, maka nilai *cluster* tersebut yang terbaik.
 - b. Metode *Silhouette*

Metode penafsiran untuk validasi *cluster* pada objekobjek. Hasil perhitungan nilai *silhouette coefficient* berada pada rentang antara -1 hingga 1. Jika nilai



silhouette semakin mendekati 1 berarti objek i sudah berada dalam *cluster* yang tepat

5. Melakukan Clustering (*K-Means Clustering* dan *Agglomerative Hierarchical Clustering*)
6. Menentukan Karakteristik dari Klaster
7. Menentukan Saran Kinerja Pemerintah Berdasarkan Klaster

LEMBAR JAWABAN FINAL 1



SEBELAS MARET STATISTICS OLYMPIAD 2020



@ihf8891p



@ssfuns



ssf.uns.ac.id



BAB II

HASIL DAN PEMBAHASAN

1. Data Collection

Sumber Data :

Naskah Final SSO SSF UNS 2020 yang bersumber dari situs resmi BPS dan kawal covid-19 diakses pada tanggal 17 Oktober 2020 16:00 WIB

Attribute Information :

Dengan informasi data adalah sebagai berikut :

Variabel	Keterangan	Jangkauan	Penjelasan
X1	Kumulatif total terinfeksi COVID-19 (jiwa)	499 - 93.356	Total orang yang positif terinfeksi COVID-19 sampai tanggal 17 Oktober 2020
X2	Kumulatif total sembuh dari COVID-19 (jiwa)	368 - 77.969	Total pasien yang dinyatakan sembuh dari COVID-19 sampai tanggal 17 Oktober 2020
X3	Total kasus positif COVID-19 aktif pada tanggal 17 Oktober 2020 (jiwa)	0 - 974	Total orang yang positif terinfeksi COVID-19 pada tanggal 17 Oktober 2020 (sudah dikurangi dengan pasien yang sembuh dan meninggal)

Variabel	Keterangan	Jangkauan	Penjelasan
X4	Total kasus sembuh dari COVID-19 pada tanggal 17 Oktober 2020 (jiwa)	0 - 1.106	Total pasien yang dinyatakan sembuh pada tanggal 17 Oktober 2020
X5	Kumulatif total meninggal akibat terinfeksi COVID-19 (jiwa)	5 - 3.529	Total orang yang meninggal akibat terinfeksi COVID-19 sampai tanggal 17 Oktober 2020
X6	Total kasus meninggal akibat COVID-19 pada tanggal 17 Oktober 2020 (jiwa)	0 - 21	Total orang yang meninggal akibat terinfeksi COVID-19 pada tanggal 17 Oktober 2020
X7	<i>Recovery Index</i> (persen)	0,38 - 0,94	Tingkat kesembuhan pasien COVID-19
X8	<i>Case Fatality Ratio</i> (persen)	0,01 - 0,07	Proporsi kematian akibat penyakit COVID-19 dibandingkan dengan





			jumlah total orang yang didiagnosis dengan penyakit tersebut selama periode tertentu yakni sampai tanggal 17 Oktober 2020
X9	Rasio lacak & isolasi (persen)	0 - 13,13	Persentase pelacakan keberadaan orang yang terinfeksi COVID-19 yang diketahui dan kemudian diisolasi
X10	Kumulatif total kasus OTG (jiwa)	6 - 385.352	Jumlah total orang dengan status OTG ("Orang Tanpa Gejala" atau istilah untuk kondisi paling ringan pada seseorang yang terinfeksi COVID-19 tapi tidak menunjukkan gejala) sampai tanggal 17 Oktober 2020

X11	Total kasus OTG selesai (jiwa)	6 - 284.304	Total orang dengan status OTG ("Orang Tanpa Gejala") pada tanggal 17 Oktober 2020
X12	Kumulatif total kasus ODP (jiwa)	646 - 96.844	Jumlah total orang dengan status ODP ("Orang Dalam Pemantauan" atau istilah untuk mengelompokkan individu berdasarkan beberapa hal seperti gejala demam atau gangguan pernapasan) sampai tanggal 17 Oktober 2020
X13	Total kasus ODP dalam proses (jiwa)	13 - 4.519	Total orang dengan status ODP ("Orang Dalam Pemantauan") dalam proses pemantauan pada tanggal 17 Oktober 2020
X14	Total kasus ODP selesai (jiwa)	735 - 92.325	Total orang dengan status ODP ("Orang Dalam Pemantauan") yang proses pemantauan telah selesai pada tanggal 17 Oktober 2020

Varia bel	Keterangan	Jangkauan	Penjelasan
X15	Kumulatif total kasus PDP (jiwa)	152 - 3.781	Jumlah total orang dengan status PDP ("Pasien Dalam





			Pemantauan” atau istilah untuk mengelompokkan individu berdasarkan beberapa hal seperti gejala demam atau gangguan pernapasan yang kondisinya semakin memburuk) sampai tanggal 17 Oktober 2020
X16	Kepadatan Penduduk (jiwa/km ²)	9 - 15.900	Perbandingan antara jumlah penduduk dan luas daerah yang ditempati atau lebih singkatnya banyaknya penduduk per satuan luas
X17	Garis Kemiskinan (Rupiah/kapita/bulan)	33.9743 - 73.2570	Tingkat minimum pendapatan yang dianggap perlu dipenuhi untuk memperoleh standar hidup yang mencukupi di suatu negara

2. Exploratory Data Analysis

Statistika Deskriptif dalam data tersebut adalah sebagai berikut :

	count	mean	std	min	25%	50%	75%	max
No	34.0	17.500000	9.958246	1.00	9.250	17.50	25.750	34.00
X1	34.0	10522.411765	17805.485760	499.00	1589.750	4052.00	11193.500	93356.00
X2	34.0	8282.117647	14885.514349	368.00	1394.500	3215.50	8111.500	77969.00
X3	34.0	126.500000	198.614023	0.00	24.500	48.50	123.250	974.00
X4	34.0	119.058824	206.392674	0.00	17.250	41.00	117.000	1106.00
X5	34.0	365.617647	701.378081	5.00	43.750	131.00	376.000	3529.00
X6	34.0	2.470588	4.560936	0.00	0.000	0.50	2.000	21.00
X7	34.0	0.758529	0.122503	0.38	0.685	0.79	0.855	0.94
X8	34.0	0.030294	0.016420	0.01	0.020	0.03	0.040	0.07
X9	34.0	2.646765	2.700350	0.00	0.710	2.20	3.390	13.13
X10	34.0	25715.529412	78780.344321	0.00	0.000	0.00	5453.500	385352.00
X11	33.0	19995.333333	63543.060652	0.00	0.000	0.00	3983.000	284304.00
X12	34.0	5430.235294	16902.228211	0.00	0.000	0.00	2551.000	96844.00
X13	34.0	294.500000	833.206521	0.00	0.000	0.00	113.000	4519.00
X14	34.0	4542.235294	16030.010348	0.00	0.000	0.00	1346.000	92325.00
X15	34.0	461.029412	885.797038	0.00	0.000	0.00	369.250	3781.00
X16	34.0	742.029412	2708.764311	9.00	54.500	103.50	261.250	15900.00
X17	34.0	457414.529412	123845.255516	0.00	395236.250	454847.00	523044.750	732570.00

3. Feature Engineering



@ihf8891p



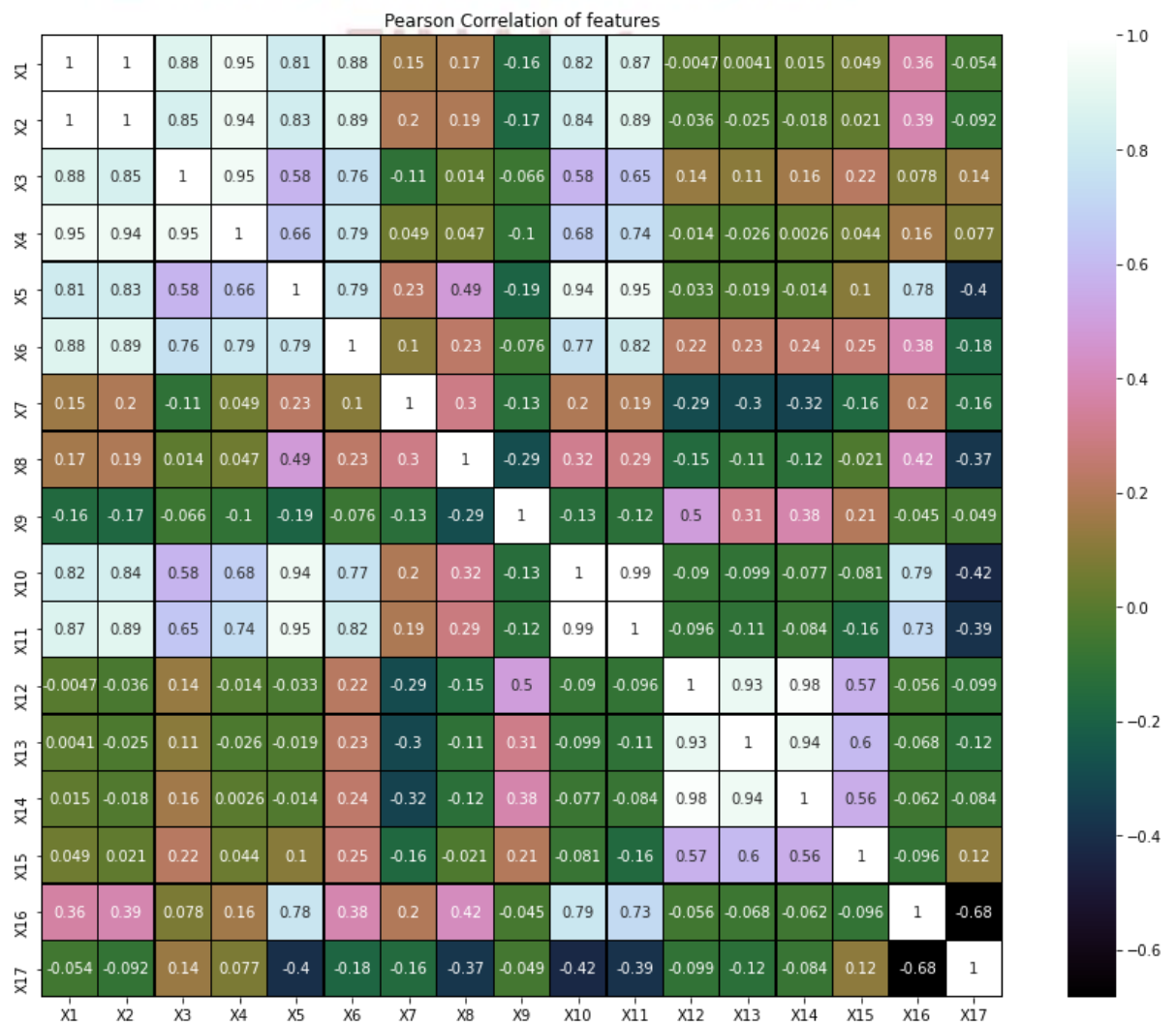
@ssfuns



ssf.uns.ac.id

Setelah ditinjau ditemukan keanehan dalam data yaitu data tidak sesuai dengan yang keterangan yang diketahui. Pada data dengan variable X10, X11, X12, X13, X14, X15, X17 ditemukan data yang bernilai 0 dan/atau NaN. Dalam keterangan data sebelumnya, Jangkauan X10, X11, X12, X13, X14, X15, X17 tidak memuat nilai 0. Diduga data tersebut merupakan data yang kosong. Untuk NaN ditemukan pada data X11 pada provinsi Jawa Tengah.

Kemudian dilakukan usaha untuk mengisi missing value tersebut. Dilakukan proses perhitungan Pearson's Correlation yang divisualisasikan dalam bentuk Heat-Map. Kemudian diperoleh grafik sebagai berikut :



Selanjutnya, ditinjau korelasi antar variable khususnya pada data yang tidak lengkap.

- X10



Dapat dilihat bahwa X10 berkorelasi tinggi dengan X2 dan X5. Kemudian dilakukan regresi linear dalam mengisi nilai kosong pada X10. Namun, setelah dilakukan regresi data yang awalnya *missing value* justru ada yang bernilai negatif. Akibatnya variable X10 tidak dapat digunakan dalam pemodelan selanjutnya.

- X11

Sama seperti X10, dapat dilihat bahwa X10 berkorelasi tinggi dengan X2 dan X5. Kemudian dilakukan regresi linear dalam mengisi nilai kosong pada X10. Namun, setelah dilakukan regresi data yang awalnya *missing value* justru ada yang bernilai negatif. Akibatnya variable X10 tidak dapat digunakan dalam pemodelan selanjutnya.

- X12, X13, X14, X15

Dapat dilihat bahwa X12, X13, X14 dan X15 berkorelasi rendah dengan variable lainnya. Sehingga tidak dilakukan proses regresi linear untuk mengisi kekosongan data. Akibatnya variabel-variabel tersebut tidak digunakan dalam pemodelan selanjutnya.

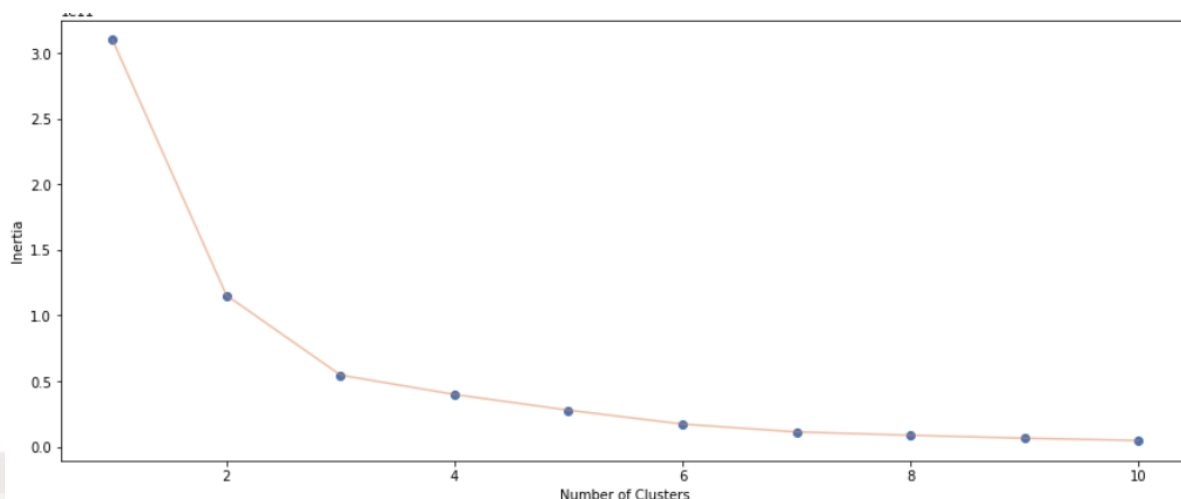
- X17

Data yang kosong adalah data Provinsi Jawa Timur. Kemudian dilakukan pengisian data X17 pada provinsi tersebut dengan nilai rata-rata dari provinsi Jawa Tengah dan DIY dengan asumsi ada keberpengaruhan angka kemiskinan dengan daerah sekitarnya.

Jadi, feature/variabel yang tidak digunakan dalam pemodelan klastering selanjutnya adalah X10, X11, X12, X13, X14, X15

4. Menentukan Jumlah Klaster

a. Metode Elbow



@ihf8891p



@ssfuns



ssf.uns.ac.id



Dari grafik di atas diperoleh bahwa jumlah klaster yang setelahnya tidak ada perubahan WCSS yang drastis adalah jumlah klaster=3.

b. Metode Silhouette

```
For n_clusters=2, The Silhouette Coefficient is 0.530659664518871
For n_clusters=3, The Silhouette Coefficient is 0.5429858339033454
For n_clusters=4, The Silhouette Coefficient is 0.4872891042142783
For n_clusters=5, The Silhouette Coefficient is 0.522886616218323
For n_clusters=6, The Silhouette Coefficient is 0.5352079262375548
For n_clusters=7, The Silhouette Coefficient is 0.5013648263382846
For n_clusters=8, The Silhouette Coefficient is 0.483223290661232
For n_clusters=9, The Silhouette Coefficient is 0.49951372221459195
For n_clusters=10, The Silhouette Coefficient is 0.48302786299627115
```

Koefisien silhouette yang paling mendekati 1 adalah pada jumlah cluster=3

Dengan demikian diperoleh jumlah cluster yang tepat berdasarkan dua metode tersebut adalah 3

5. Melakukan Clustering

1. K-Means clustering

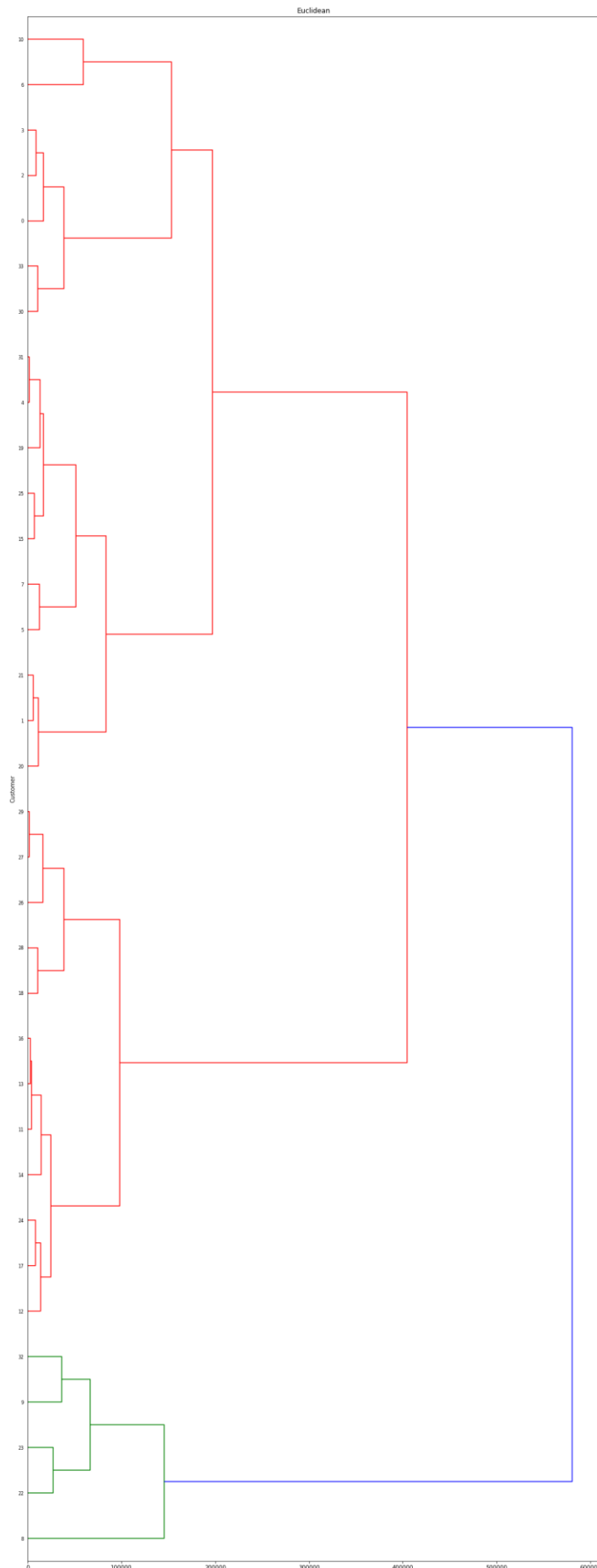
(Feature lebih dari 3 tidak bisa divisualisasikan)

2. Agglomerative Hierarchical Clustering Menentukan Karakteristik dari Klaster





SSF
SEBELAS MARET STATISTICS FAIR
2020



VABAN

1

STATISTICS
AD



@ihf8891p



@ssfuns



ssf.uns.ac.id



Diperoleh kluster

Provinsi Asal	Klaster
Jawa Barat	0
Jawa Tengah	0
Sulawesi Tenggara	0
Sulawesi Selatan	0
Nusa Tenggara Timur	0
DI Yogyakarta	1
Jambi	1
Kalimantan Barat	1
Kalimantan Timur	1
Kalimantan Tengah	1
Kalimantan Selatan	1
Kepulauan Riau	1
Nusa Tenggara Barat	1
Sumatera Selatan	1
Sulawesi Tengah	1
Riau	1
Maluku Utara	1
Maluku	1
Papua Barat	1
Aceh	2
Bali	2
Banten	2
Bangka Belitung	2
Bengkulu	2
DKI Jakarta	2
Jawa Timur	2
Kalimantan Utara	2
Sumatera Barat	2
Sulawesi Utara	2
Sumatera Utara	2
Lampung	2
Papua	2
Sulawesi Barat	2
Gorontalo	2

6. Menentukan Karakteristik dari Klaster

No	Provinsi Asal	X1	X2	X5	X7	clusters
----	---------------	----	----	----	----	----------



@ihf8891p



@ssfuns



ssf.uns.ac.id



9	Jawa Barat	30043,000000	19924,000000	559,000000	0.66	0
10	Jawa Tengah	28723,000000	22980,000000	1579,000000	0.80	0
23	Sulawesi Tenggara	4023,000000	2517,000000	71,000000	0.63	0
24	Sulawesi Selatan	17352,000000	14828,000000	444,000000	0.86	0
33	Nusa Tenggara Timur	584,000000	406,000000	7,000000	0.70	0
6	DI Yogyakarta	3285,000000	2551,000000	85,000000	0.78	1
8	Jambi	968,000000	368,000000	18,000000	0.38	1
12	Kalimantan Barat	1413,000000	1224,000000	9,000000	0.87	1
13	Kalimantan Timur	11503,000000	8541,000000	435,000000	0.74	1
14	Kalimantan Tengah	4081,000000	3437,000000	145,000000	0.84	1
15	Kalimantan Selatan	11400,000000	9984,000000	451,000000	0.88	1
17	Kepulauan Riau	2768,000000	2246,000000	68,000000	0.81	1
18	Nusa Tenggara Barat	3673,000000	2994,000000	209,000000	0.82	1
19	Sumatera Selatan	7104,000000	5381,000000	387,000000	0.76	1
25	Sulawesi Tengah	668,000000	436,000000	25,000000	0.65	1
27	Riau	11358,000000	6823,000000	259,000000	0.60	1
28	Maluku Utara	2120,000000	1906,000000	74,000000	0.90	1
29	Maluku	3521,000000	2651,000000	42,000000	0.75	1
30	Papua Barat	3465,000000	2504,000000	46,000000	0.72	1
1	Aceh	6436,000000	4318,000000	226,000000	0.67	2
2	Bali	10697,000000	9505,000000	343,000000	0.89	2
3	Banten	7593,000000	5735,000000	226,000000	0.76	2
4	Bangka Belitung	499,000000	430,000000	6,000000	0.86	2
5	Bengkulu	862,000000	701,000000	43,000000	0.81	2
7	DKI Jakarta	93356,000000	77969,000000	2015,000000	0.84	2
11	Jawa Timur	48690,000000	42455,000000	3529,000000	0.87	2
16	Kalimantan Utara	691,000000	597,000000	5,000000	0.86	2
20	Sumatera Barat	10700,000000	5756,000000	187,000000	0.54	2
21	Sulawesi Utara	4959,000000	4142,000000	185,000000	0.84	2
22	Sumatera Utara	11861,000000	9497,000000	497,000000	0.80	2



@ihf8891p



@ssfun



ssf.uns.ac.id



26	Lampung	1285,000000	869,000000	48,000000	0.68	2
31	Papua	8168,000000	4438,000000	117,000000	0.54	2
32	Sulawesi Barat	949,000000	704,000000	12,000000	0.74	2
34	Gorontalo	2964,000000	2775,000000	79,000000	0.94	2

Pada Variabel X1 : Kumulatif total terinfeksi COVID-19 (jiwa)

Pada Variabel X2 : Kumulatif total sembuh dari COVID-19 (jiwa)

Pada Variabel X5 : Kumulatif total meninggal akibat terinfeksi COVID-19 (Jiwa)

Pada Variabel X7 : Recovery Index

Statistika deskriptif untuk setiap klaster dan setiap variable diperoleh sebagai karakteristik setiap klaster

7. Menentukan Saran Kinerja Pemerintah Berdasarkan Klaster

Klaster 1

Klaster 1 memiliki karakteristik dengan variabel jumlah kasus aktif dan jumlah kasus kematian/juta penduduk yang tinggi. Oleh karena itu kedua variabel ini menjadi prioritas oleh pemerintah untuk dilakukan berbagai kebijakan.

Beberapa kebijakan untuk menurunkan jumlah kasus aktif antara lain:

- Membatasi akses keluar masuk wilayah.
- Membatasi aktivitas masyarakat dengan cara belajar, bekerja, dan beribadah di dalam rumah.
- Menutup tempat-tempat pusat keramaian seperti pasar dan *mall*.
- Memperketat pengawasan terhadap masyarakat dalam menjalankan protokol kesehatan.
- Memberikan sanksi yang tegas kepada setiap masyarakat yang melanggar peraturan pemerintah dalam penanganan penyebaran COVID-19.
- Menyediakan cairan disinfektan di tempat-tempat publik.
- Menerapkan pembatasan fisik dan sosial di tempat-tempat publik.





h. Memperbanyak jumlah tes cepat dan tes *polymerase chain reaction* (PCR) secara massal.

i. Mewajibkan untuk memakai masker di ruang publik atau luar rumah.

j. Melakukan penelusuran kontak dari kasus positif secara cepat dan tanggap.

k. Melakukan isolasi bagi yang terinfeksi COVID-19.

l. Memberikan edukasi-edukasi kepada masyarakat mengenai cara memutus mata rantai penyebaran COVID-19 seperti sering mencuci tangan dengan air mengalir dan sabun selama 20 detik, melakukan pembatasan fisik dan sosial, tidak menyentuh area wajah, mencuci bahan makanan, menerapkan etika saat batuk dan bersin, tidak berbagi barang pribadi, membersihkan perabotan rumah, dan menjaga imunitas tubuh.

Beberapa kebijakan untuk menurunkan jumlah kasus kematian/juta penduduk antara lain:

a. Menjamin tersedianya tenaga medis yang memadai dalam menangani para pasien di setiap tingkat daerah sesuai jumlah penduduk di daerah tersebut.

b. Memperbanyak fasilitas pelayanan kesehatan khusus dalam penanganan COVID-19 di setiap tingkat daerah sesuai jumlah penduduk di daerah tersebut.

c. Melakukan pengawasan dan kontrol yang lebih ketat di setiap tingkat daerah.

d. Melakukan koordinasi yang lebih baik antara pemerintah pusat dengan kepala-kepala tingkat daerah.

e. Menjamin ketersediaan alat pelindung



@ihf8891p



@ssfuns



ssf.uns.ac.id



diri (APD) bagi tenaga medis di setiap fasilitas kesehatan.

f. Menangani pasien darurat yang lebih responsif dan sesuai dengan prosedur kesehatan.

g. Melakukan isolasi mandiri bagi orang yang berisiko tinggi terinfeksi COVID-19 seperti lansia yang memiliki penyakit penyerta.

h. Memberikan masker gratis kepada seluruh penduduk miskin.

i. Memberikan edukasi kepada masyarakat untuk segera memeriksakan diri ke rumah sakit apabila mengalami gejala-gejala COVID-19.

2. Klaster 2

Klaster 2 memiliki karakteristik dengan variabel jumlah kasus kematian tertinggi. Oleh karena variabel ini menjadi prioritas oleh pemerintah untuk dilakukan berbagai kebijakan.

Beberapa kebijakan pemerintah untuk menurunkan jumlah kasus kematian antara lain:

a. Menjamin tersedianya tenaga medis yang memadai dalam menangani para pasien.

b. Memperbanyak fasilitas pelayanan kesehatan khusus dalam penanganan COVID-19.

c. Menjamin ketersediaan alat pelindung diri (APD) bagi tenaga medis di setiap fasilitas kesehatan.

d. Menangani pasien darurat yang lebih responsif dan sesuai dengan prosedur kesehatan.

e. Melakukan isolasi mandiri bagi orang yang berisiko tinggi terinfeksi COVID-19



@ihf8891p



@ssfuns



ssf.uns.ac.id



seperti lansia yang memiliki penyakit penyerta.

f. Memberikan masker gratis kepada masyarakat miskin.

g. Memberikan edukasi kepada masyarakat untuk segera memeriksakan diri ke rumah sakit apabila mengalami gejala-gejala COVID-19.

3. Klaster 3

Klaster 3 memiliki karakteristik dengan variabel jumlah kasus aktif pada kategori sedang. Oleh karena itu variabel ini menjadi prioritas untuk direkomendasikan berbagai kebijakan kepada pemerintah. Beberapa kebijakan untuk menurunkan jumlah kasus aktif antara lain:

- a. Merekomendasikan untuk menutup tempat-tempat pusat keramaian seperti pasar dan *mall*.
- b. Merekomendasikan untuk menyediakan cairan disinfektan di tempat-tempat publik.
- c. Merekomendasikan untuk memperbanyak jumlah tes cepat dan tes *polymerase chain reaction* (PCR) secara massal.
- d. Merekomendasikan untuk menerapkan pembatasan fisik dan sosial di tempat-tempat publik.
- e. Merekomendasikan untuk melakukan penelusuran kontak dari kas





LEMBAR JAWABAN FINAL 1

BAB III **KESIMPULAN**

1. Kelompok/cluster provinsi di Indonesia berdasarkan data dibagi menjadi 3 kelompok/cluster
2. Saran kinerja pemerintah dalam penanganan COVID-19 yang lebih efektif dan efisien berdasarkan karakteristik dari setiap klaster berdasarkan analisis sebelumnya

