

# PENGELOMPOKAN PROVINSI TERDAMPAK *CORONA VIRUS DISEASE-19* (*COVID-19*) BERDASARKAN INDIKATOR KESEHATAN LINGKUNGAN DI INDONESIA

Maysarah Sabariah Kudadiri (G14170084)

Dosen Pembimbing :

1. Pika Silvianti, S.Si, M.Si
2. Farit Mochamad Afendi, S.Si M.Si



# POKOK BAHASAN

**Latar  
Belakang**

**Tinjauan  
Pustaka**

**Metodologi  
Penelitian**

**Daftar  
Pustaka**

# Latar Belakang

# Latar Belakang



Indonesia melaporkan pada tahun 2020 terdapat 743.198 kasus Covid-19.



*Hygiene* sanitasi bertujuan pencegahan dan pemutusan mata rantai penularan penyakit

# Latar Belakang

Indikator Kesehatan Lingkungan Tahun 2020



Analisis Gerombol



*Cluster Ensemble*  
(Strehl dan Ghosh 2003)

# Tujuan Penelitian

Mengelompokkan provinsi terdampak *Corona Virus Disease-19* (COVID-19) berdasarkan indikator kesehatan lingkungan agar mempermudah pemerintah dalam menentukan prioritas pembangunan kesehatan lingkungan di Indonesia tahun 2020

# Tinjauan Pustaka

## Analisis Gerombol

Analisis gerombol adalah metode membagi  $n$  objek menjadi  $m$  gerombol ( $m < n$ ) berdasarkan karakteristiknya (Johnson dan Wichern 2007).

Jarak Euclidean digunakan sebagai ukuran kemiripan atau ketidakmiripan antar objek

Jarak *Euclidean* :

$$d_{kl} = \|x_{kj} - x_{lj}\| = \sqrt{\sum_{j=1}^p (x_{kj} - x_{lj})^2}$$

$d_{kl}$  : *jarak antara objek ke – k dengan objek ke – l*

$x_{kj}$  : *data ke – k pada peubah ke – j*

$x_{lj}$  : *data ke – l pada peubah ke – j*

Metode gerombol dibagi menjadi :

1. Metode Hierarki
2. Metode tak berhierarki



# Metode Hierarki

Metode berhierarki digunakan jika peneliti belum mengetahui banyaknya gerombol yang akan dibentuk dan akan menggerombolkan objek secara tersrtuktur (Mattjik & Sumertajaya 2011).

## *Agglomerative hieararchical clustering*

Dalam metode *agglomerative hieararchical clustering* terdapat lima metode perbaikan jarak yang dapat digunakan, diantaranya:

# Metode Hierarki

1

## PAUTAN TUNGGA

$$d_{(uv)w} = \min(d_{uw}, d_{vw})$$

2

## PAUTAN LENGKAP

$$d_{(uv)w} = \max(d_{uw}, d_{vw})$$

3

## PAUTAN CENTROID

$$d_{(uv)w} = \frac{n_u d_{uw} + n_v d_{vw}}{n_u + n_v} - \frac{n_u n_v d_{uv}}{(n_u + n_v)^2}$$

4

## PAUTAN RATAAN

$$d_{(uv)w} = \frac{\sum i \sum k d_{ik}}{n_{uv} n_w}$$

5

## PAUTAN WARD

$$SSE = \sum_{j=1}^p (\sum_{i=1}^n x_{ij}^2 - \frac{1}{n} (\sum_{i=1}^n x_{ij})^2)$$

# Metode Tak Berhierarki

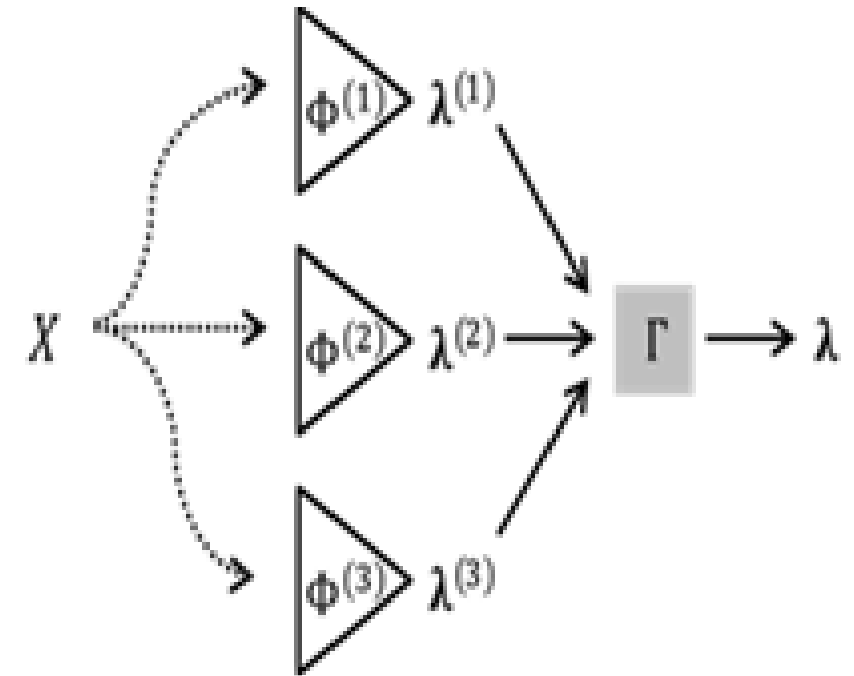
Metode tak berhierarki digunakan untuk mengelompokkan objek dimana banyaknya gerombol yang akan dibentuk dapat ditentukan terlebih dahulu. Metode tak berhierarki yang umum digunakan adalah k-rataan (Belinda *et al.* 2019)

## Langkah-langkah metode k-rataan:

- 01 Bagi objek ke dalam k gerombol awal
- 02 Masukkan tiap objek ke suatu gerombol berdasarkan rataan terdekat
- 03 Ulangi langkah 2 sampai tidak ada lagi pemindahan

# *Cluster Ensemble*

*Cluster Ensemble* diperkenalkan oleh Strehl dan Gosh (2003), yaitu sebuah metode yang digunakan untuk mengombinasikan sekumpulan solusi gerombol. Metode ini memiliki keunggulan dibandingkan dengan metode penggerombolan lainnya yaitu dapat meningkatkan kualitas dan kekekaran solusi gerombol.



Gambar 1 Kerangka ilustrasi *cluster ensemble*

## Nilai *Reproducibility*

Nilai *reproducibility* digunakan untuk membandingkan solusi gerombol .

### Langkah-langkah menghitung nilai *reproducibility*:

- 01 Membuat tabulasi silang solusi gerombol ke-i dengan solusi gerombol ke-j untuk  $i = j$ .
- 02 Selanjutnya dihitung  $R_{ij}$  (jumlah diagonal/jumlah objek)
- 03 Hitung nilai *reproducibility* total solusi gerombol

## Evaluasi Hasil Gerombol

Pengukuran validitas yang dapat digunakan adalah compactness (CP). CP mengukur jarak rata-rata antara setiap pasang titik data yang terdapat di dalamnya gerombol yang sama.

Rumus :

$$CP = \frac{1}{N} \sum_{k=1}^K n_k \left( \frac{\sum_{x_i, x_j \in C_k} d(x_i, x_j)}{n_k(n_k-1)/2} \right)$$

Keterangan

- $K$  : Banyaknya gerombol yang terbentuk
- $n_k$  : Banyaknya objek yang termasuk kedalam gerombol ke- $k$
- $d(x_i, x_j)$  : Jarak antara objek ke- $i$  dan objek ke- $j$
- $N$  : Banyaknya seluruh objek

# Metode Penelitian

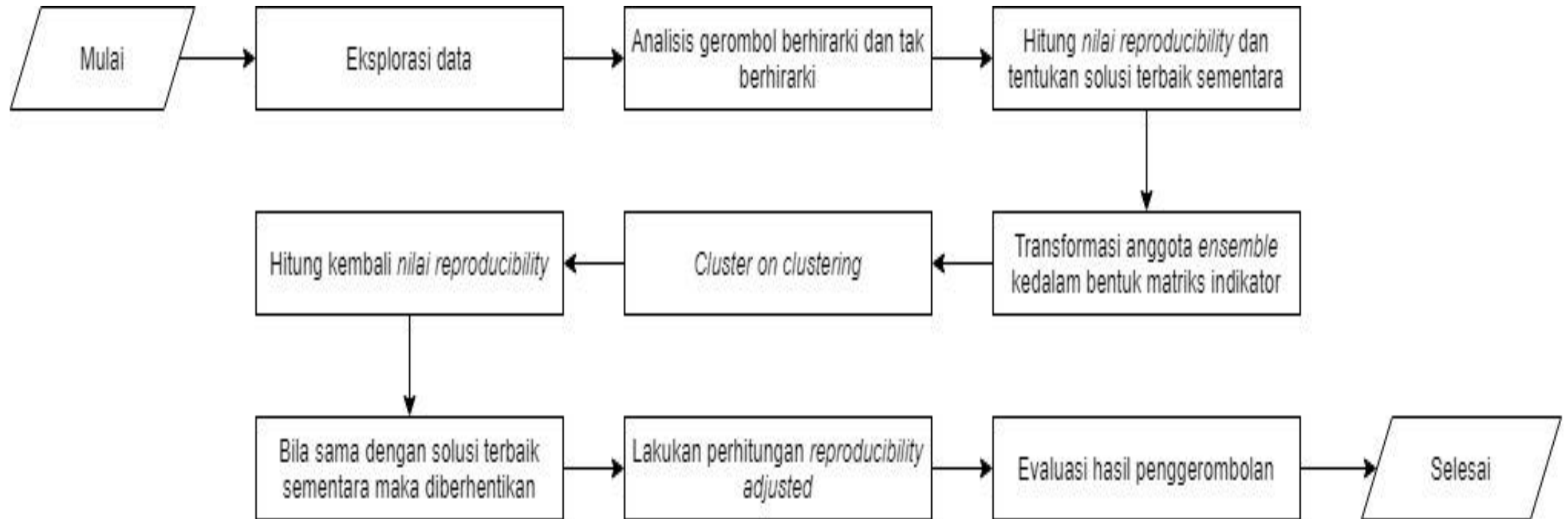
# Data

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari Kementerian Kesehatan Republik Indonesia, yaitu data indikator kesehatan lingkungan dari seluruh provinsi Indonesia tahun 2020. Meliputi 34 provinsi dan 6 peubah berskala numerik

Kode	Peubah
X1	Pemenuhan kualitas kesehatan lingkungan (%)
X2	Pelaksanaan sanitasi total berbasis masyarakat (%)
X3	Pengawasan kualitas air minum (%)
X4	Pengelolaan limbah medis (%)
X5	Tempat Pengolahan makanan memenuhi syarat kesehatan (%)
X6	Tatanan kawasan sehat (%)



# Prosedur Analisis Data



# Daftar Pustaka

- Belinda NS, HG IR, Yozza H. 2019. Penerapan analisis cluster ensemble dengan metode rock untuk mengelompokkan provinsi di indonesia berdasarkan indikator kesejahteraan rakyat. *J. Mat. UNAND*. 8 (2): 108-119
- Covid19 (2020) Peta Sebaran. Tersedia pada: <https://covid19.go.id/peta-sebaran> (Diakses: 11 Agustus 2020).
- Johnson RA, Winchern DW. 2007. Applied Multivariate Statistical Analysis (sixth ed). Prentice Hall, New Jersey.
- Hu X, Yoo I. 2004. Cluster Ensemble and Its Applications in Gene Expression Analysis BT - Second Asia-Pacific Bioinformatics Conference (APBC2004). 29:297–302.
- Mattjik AA, Sumertajaya IM. 2011. Sidik Peubah Ganda dengan Menggunakan SAS. Wibawa GNA, Hadi AF, editor. Bogor (ID): IPB Press.
- Mayasari TR. 2020. Pengelompokan provinsi berdasarkan variabel kesehatan lingkungan dan pengaruhnya terhadap kemiskinan di indonesia tahun 2018. *J. SigerMat*. 1 (1):2430-2471.
- Orme, B. & Johnson, R. 2008. *Improving K-means Cluster Analysis : Ensemble Analysis Instead Of Highest Reproducibility Replicates*. Sawtooth Software.
- Putri YD, HG IR, Yozza H. 2019. Penerapan Analisis Cluster Ensemble Untuk Mengelompokkan provinsi Di Indonesia berdasarkan indikator kesehatan lingkungan. *J. Mat. UNAND*. 8 (1): 323-331

- Rembulan GD, Wijaya T, Palullungan D, Alfina KN, Qurthuby M. 2020. Kebijakan pemerintah mengenai coronavirus disease (COVID-19) di setiap provinsi di indonesia berdasarkan analisis klaster. *JIEMS Journal Ind. Eng. Manage. Syst.* 13 (2)
- Strehl A, Gosh J. 2003. A knowledge reuse framework for combining partitionings. *The Journal of Machine learning Research.* 3 (3): 583- 617.
- WHO, UNICEF. 2020. Air, sanitasi, higiene, dan pengelolaan limbah yang tepat dalam penanganan wabah COVID-19. *World Heal. Organ.*:1–10.

# TERIMA KASIH

