

# Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Ketenagakerjaan Menggunakan Metode Ward

Arbi Wijaya<sup>1</sup>, Helma<sup>2</sup>

<sup>12</sup>Program Studi Matematika, Universitas Negeri Padang  
e-mail: [arbiwijaya8@gmail.com](mailto:arbiwijaya8@gmail.com)

## Abstrak

Penelitian ini bertujuan untuk mengetahui hasil pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat berdasarkan indikator Ketenagakerjaan dengan menggunakan metode Ward. Penelitian ini merupakan penelitian terapan dan jenis data yang digunakan adalah data sekunder tentang ketenagakerjaan di Provinsi Sumatera Barat tahun 2022. Sumber data diperoleh dari website resmi Badan Pusat Statistika Sumatera Barat yaitu <https://sumbar.bps.go.id/>. Dengan menggunakan 4 variabel yaitu Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), Tingkat Kesempatan Kerja (TKK) dan Tingkat Pekerja Tidak Penuh (TPTP). Pada penelitian ini dilakukan proses clustering dengan menggunakan jarak Square Euclidean dan validasi cluster dengan menggunakan Dunn Index. Hasil penelitian menunjukkan bahwa cluster yang terbentuk sebanyak 3 cluster, yaitu cluster 1 sebanyak 1 Kabupaten/Kota termasuk tingkatan ketenagakerjaan tinggi, cluster 2 sebanyak 12 Kabupaten/Kota termasuk tingkatan ketenagakerjaan sedang, cluster 3 sebanyak 6 Kabupaten/Kota termasuk tingkatan ketenagakerjaan rendah. Hasil cluster berdasarkan nilai validasi cluster Dunn Index sebesar 0.780.

**Kata kunci:** *Ketenagakerjaan, Analisis Cluster, Ward*

## Abstract

This research aims to determine the results of the grouping of districts/cities in West Sumatra Province based on employment indicators using the Ward method. This research is applied research and the type of data used is secondary data about employment in West Sumatra Province in 2022. The data source was obtained from the official website of the West Sumatra Central Statistics Agency, namely <https://sumbar.bps.go.id/>. By using 4 variables, namely the Open Unemployment Rate (TPT), Labor Force Participation Rate (TPAK), Employment Opportunity Rate (TKK) and Partial Employment Rate (TPTP). In this research, a clustering process was carried out using Square Euclidean distance and cluster validation using the Dunn Index. The results of the research show that there are 3 clusters formed, namely cluster 1 with 1 Regency/City including high employment levels, cluster 2 with 12 Regency/Cities including medium employment levels, cluster 3 with 6 Regency/Cities including low employment levels. The cluster results are based on the Dunn Index cluster validation value of 0.780.

**Keywords :** *Employment, Cluster Analysis, Ward*

## PENDAHULUAN

Indonesia merupakan salah satu negara berkembang yang sedang meningkatkan perekonomian nasional dikarenakan banyaknya permasalahan dalam kesejahteraan masyarakat. Salah satu sektor yang sangat penting dalam kesejahteraan masyarakat dalam membangun perekonomian nasional adalah

Ketenagakerjaan. Ketenagakerjaan merupakan segala sesuatu yang memiliki keterkaitan dengan tenaga kerja baik itu pada waktu sebelumnya, selama dan setelah selesainya masa hubungan kerja. Berdasarkan Badan Pusat Statistik (BPS), tingkat ketenagakerjaan suatu Kabupaten/Kota dapat dilihat melalui beberapa kategori diantaranya yaitu Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), Tingkat Kesempatan Kerja (TKK) dan Tingkat Pekerja Tidak Penuh (TPTP).

TPT di Provinsi Sumatera Barat mengalami penurunan sebesar 0,24 persen poin dari 6,52 persen pada Agustus 2021 menjadi 6,28 persen pada Agustus 2022. Di level nasional, TPT antar provinsi di Indonesia menunjukkan angka yang cukup bervariasi. Provinsi Sumatera Barat termasuk dalam 10 Provinsi dengan TPT tertinggi di Indonesia dengan posisi ketujuh dan tertinggi kedua di Pulau Sumatera setelah Provinsi Kepulauan Riau. Pada tahun 2018-2022, TPT di Sumatera Barat mengalami fase naik turun. Pada tahun 2022, Kota Padang memiliki TPT tertinggi sebesar 11,69 persen dan Kabupaten Kepulauan Mentawai memiliki TPT terendah sebesar 1,39 persen (BPS Sumbar, 2023).

Tingkat ketenagakerjaan di Provinsi Sumatera Barat masih dikatakan belum termasuk tingkat ketenagakerjaan yang tinggi, hal ini dapat dilihat dari persebaran jumlah tenaga kerja di setiap Kabupaten/Kota yang ada di Provinsi Sumatera Barat yang pada saat ini masih belum merata. Kebijakan ketenagakerjaan diharapkan dapat menyelesaikan permasalahan kesenjangan antar wilayah Kabupaten/Kota di Provinsi Sumatera Barat. Dalam meningkatkan ketenagakerjaan diperlukan identifikasi Kabupaten/Kota mana saja yang memiliki tingkat ketenagakerjaan yang tinggi maupun rendah. Berdasarkan hal tersebut, maka perlu dilakukan pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat menjadi beberapa kelompok berdasarkan indikator ketenagakerjaan.

Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat berdasarkan indikator ketenagakerjaan dapat dilakukan dengan analisis *cluster*. Analisis *cluster* merupakan salah satu analisis multivariate yang bertujuan untuk mengelompokan data berdasarkan karakteristik yang serupa (Nugroho, 2008). Analisis *cluster* merupakan analisis pengelompokan objek-objek yang sejenis menjadi satu kelompok yang sama, dengan ciri antar anggota *cluster* memiliki homogenitas (kesamaan) yang tinggi dalam satu *cluster* (*within cluster*) dan heterogenitas (perbedaan) antar objek dalam *cluster* lain (*between cluster*) (Han, Kamber & Pei, 2012). Analisis *cluster* terdapat dua metode yaitu metode hierarki dan non hierarki. Pada metode hierarki terdiri dari dua bagian yaitu *agglomerative* dan *divisive*.

Metode ward dikembangkan oleh Joe H. Ward Jr. pada tahun 1963. Metode ward merupakan metode pembentukan *cluster* berdasarkan pada hilangnya informasi akibat menggabungkan objek menjadi *cluster*. Hal tersebut diukur dengan menggunakan jumlah total dari deviasi kuadrat pada mean *cluster* dalam setiap pengamatan. Ukuran yang digunakan adalah *Error sum of squares* (SSE) (Johnson & Wichern, 2007). Alasan menggunakan metode ward karena lebih efisien dan mampu membuat sebuah *cluster* hingga ukuran sangat kecil (Govender & Sivakumar, 2019).

Dalam proses analisis *cluster*, dilakukan pengukuran antar data menggunakan jarak *square euclidean*. Jarak *squared euclidean* merupakan variasi dari jarak *euclidean*, dimana dalam jarak *squared euclidean* menghapus akar pada persamaan kuadratnya. Ukuran jarak antar data berguna untuk mengetahui kemiripan antara dua objek dengan melihat jarak yang paling terdekat, karena semakin kecil jarak antara dua objek maka akan semakin mirip pula karakteristiknya. Adapun persamaan jarak *square euclidean* sebagai berikut:

$$d_{i,j} = \sum_{k=1}^p (x_{ik} - x_{jk})^2 \quad (1)$$

Langkah selanjutnya yaitu memperbarui matriks jarak untuk mencari *cluster* terdekat dari 18 *cluster*. Jarak antar *cluster* dihitung berdasarkan metode *ward* dengan mencari nilai *SSE* terkecil. Adapun persamaan *SSE* sebagai berikut:

$$SSE_{i,j} = \frac{n_A n_B}{n_A + n_B} ((\bar{y}_A - \bar{y}_B)'(\bar{y}_A - \bar{y}_B)) \quad (2)$$

Hal utama dari analisis *cluster* adalah menentukan banyak *cluster*. Menurut Hair (2010), menentukan banyaknya *cluster* yang paling umum dan sederhana yaitu dengan cara melihat perubahan persentase pada kemiripan dalam *cluster*. Misalnya dengan menggunakan nilai *coefficient* pada *agglomerative* dalam *software SPSS*, namun untuk metode *word* ukuran jarak yang digunakan yaitu jumlah kuadrat dalam *cluster*. Banyaknya *cluster* yang terbentuk adalah hasil pengurangan dari banyaknya objek yang diamati ( $n$ ) dengan *stage* saat proses dihentikan. Adapun banyaknya *stage* yaitu  $n-1$ .

Selanjutnya, tahap interpretasi meliputi pemberian nama atau keterangan pada masing-masing *cluster* yang terbentuk sebagai gambaran sifat dari *cluster* yang menjelaskan bagaimana antar *cluster* dapat menggunakan nilai rata-rata setiap *cluster* pada setiap variabel. data yang digunakan adalah data awal penelitian (Hair, 2010).

Proses terakhir dalam analisis *cluster* yaitu menguji kevalidan hasil *cluster*. Pengujian dilakukan dengan metode *Dunn Index*. *Dunn Index* pertama kali dikenalkan oleh J.C Dunn pada tahun 1973. *Dunn index* merupakan ukuran validasi *cluster* yang berlandaskan pada fakta bahwa *cluster* yang terpisah itu memiliki jarak antar *cluster* yang besar dan jarak antar data dalam *cluster* yang kecil. *Dunn Index* memiliki rentang nilai dari 0 sampai tak hingga. Jumlah *cluster* yang terbaik jika nilai memiliki *Dunn Index* yang tertinggi (Brock, 2008). Adapun persamaan jarak *square euclidean* sebagai berikut:

$$DI = \frac{\min_{i=1+1, \dots, N}(d(c_i, c_j))}{\max_{m=1, \dots, k}(diam(c_k))} \quad (3)$$

## METODE

Penelitian ini merupakan penelitian dasar/teori yang membahas tentang pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat berdasarkan indikator ketenagakerjaan. Metode yang digunakan dalam penelitian ini adalah metode *ward*.

Berikut ini langkah-langkah yang dilakukan:

1. Mengumpulkan literatur mengenai analisis *cluster*.
2. Mengambil data berdasarkan indikator ketenagakerjaan yang diperoleh dari website resmi Badan Pusat Statistika Sumatera Barat yaitu <https://sumbar.bps.go.id>.
3. Melakukan deskripsi data terhadap variabel yang digunakan.
4. Melakukan proses standarisasi data untuk menstandarkan satuan data agar pada proses *cluster* tidak terjadi masalah. Jika satuan data sudah setara tidak perlu dilakukan proses standarisasi data. Lanjut ke langkah selanjutnya yaitu proses analisis *cluster*.
5. Langkah pertama dalam analisis *cluster* yaitu menghitung jarak antar dua objek menggunakan persamaan (1).
6. Melakukan proses analisis *cluster hierarki* menggunakan metode *ward*. Adapun langkah-langkah dalam metode *ward* sebagai berikut:
  - a. Menghitung nilai *SSE* pada setiap objek data satu-persatu dengan persamaan (2).
  - b. Setelah nilai *SSE* diketahui, maka akan dibuat dalam bentuk matriks *D*. Proses ini dinamakan tahapan pembentukan *cluster*.
  - c. Membentuk *cluster* berdasarkan tahapan pembentukan *cluster* dengan nilai *SSE* yang terkecil adalah sebuah *cluster*.

- d. Memberi nama pada setiap *cluster* yang sudah terbentuk. Hasil dari metode *ward* adalah bentuk diagram berupa *dendrogram*.
7. Menentukan jumlah *cluster* yang terbentuk dengan melihat perubahan persentase yang tertinggi.
8. Melakukan interpretasi *profiling cluster*.
9. Validasi hasil *cluster* menggunakan *Dunn Index* dengan Persamaan (3).
10. Menarik kesimpulan.

## HASIL DAN PEMBAHASAN

### Analisis Deskriptif

Dalam pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat menggunakan Metode Ward dengan indikator ketenagakerjaan yaitu TPT, TPAK, TKK, TPP. Pengolahan analisis cluster untuk pengelompokan Kabupaten/Kota dengan Metode Ward menggunakan Software Microsoft Excel dan pengujian hasilnya menggunakan Software SPSS. Adapun gambaran tentang indikator ketenagakerjaan di Provinsi Sumatera Barat dapat dilihat pada Tabel 1.

**Tabel 1. Statistik Deskriptif**

Variable	N	Min	Max	Mean	Std. Deviation
TPT ( $X_1$ )	19	1.39	11.69	5.28	1.95
TPAK ( $X_2$ )	19	62.81	78.25	70.24	4.62
TKK ( $X_3$ )	19	88.31	98.61	94.72	1.95
TPTP ( $X_4$ )	19	19.03	58.26	33.43	10.09

Dari Tabel 1, dapat diketahui bahwa variabel  $X_1$  memiliki nilai minimum sebesar 1.39 dan nilai maksimum sebesar 11.69 dengan nilai rata-rata  $X_1$  sebesar 5,28 dan nilai standar deviasi sebesar 1.95. variabel  $X_2$  memiliki nilai minimum sebesar 62.81 dan nilai maksimum sebesar 78.25 dengan nilai rata-rata  $X_2$  sebesar 70.24 dan nilai standar deviasi sebesar 4.62. variabel  $X_3$  memiliki nilai minimum sebesar 88.31 dan nilai maksimum sebesar 98.61 dengan nilai rata-rata  $X_3$  sebesar 98.61 dan nilai standar deviasi sebesar 1.95. variabel  $X_4$  memiliki nilai minimum sebesar 19.03 dan nilai maksimum sebesar 58.26 dengan nilai rata-rata  $X_4$  sebesar 33.43 dan nilai standar deviasi sebesar 10.09.

### Hasil Analisis Pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat berdasarkan indicator ketenagakerjaan metode Word

Dalam proses analisis *cluster*, dilakukan perhitungan ukuran antar data menggunakan jarak *square euclidean* yang berguna untuk mengetahui kemiripan antara dua objek dengan melihat jarak yang paling terdekat, karena semakin kecil jarak antara dua objek maka akan semakin mirip pula karakteristiknya. Adapun perhitungan jarak *square euclidean* antara objek pertama dengan objek kedua menggunakan persamaan (1) sebagai berikut:

$$\begin{aligned}
 d_{(1,2)} &= ((1.39 - 4.61)^2 + (78.25 - 66.95)^2 + (98.61 - 95.39)^2 + (58.26 - 42.01)^2) \\
 &= ((-3.22)^2 + (11.3)^2 + (3.22)^2 + (16.25)^2) \\
 &= (10.3684 + 127.69 + 10.3684 + 264.0625) \\
 &= 412.4893
 \end{aligned}$$

Demikian seterusnya dengan langkah yang sama untuk perhitungan jarak *square euclidean* pada semua objek, sehingga diperoleh hasil jarak terkecilnya. Pada penelitian ini pasangan *cluster* 16 (Kota Padang Panjang) dan *cluster* 17 (Kota Bukittinggi) memiliki ukuran jarak yang terkecil yaitu sebesar 0.254. Maka, sekarang telah terbentuk 18 *cluster*.

Langkah selanjutnya yaitu memperbarui matriks jarak untuk mencari *cluster* terdekat dari 18 *cluster*. Jarak antar *cluster* dihitung berdasarkan metode *ward* dengan mencari nilai *SSE* terkecil. Adapun perhitungan nilai *SSE* pada *cluster* 16 (Kota

Padang Panjang) dan *cluster* 17 (Kota Bukittinggi) dapat menggunakan persamaan (2) sebagai berikut:

$$SSE_{(16,17)} = \frac{n_{16}n_{17}}{n_{16} + n_{17}} ((\bar{y}_{16} - \bar{y}_{17})'(\bar{y}_{16} - \bar{y}_{17}))$$

$$SSE_{(16,17)} = \frac{(1)(1)}{(1) + (1)} (0.254)$$

$$SSE_{(16,17)} = 0.127$$

Demikian seterusnya dengan langkah yang sama untuk perhitungan varians pada semua *cluster* hingga semua menjadi satu *cluster*. Proses ini dinamakan proses *agglomerasi* yang dimana proses ini mengelompokan dua atau lebih objek yang terdekat menjadi 1 *cluster*, kemudian menghitung jarak suatu *cluster* dengan objek yang baru lagi secara bertahap satu per satu. Dengan menggunakan bantuan *Software* SPSS, maka diperoleh *Agglomeration Schedule* menggunakan metode Ward yang dapat dilihat pada Tabel 2.

**Tabel 2. Agglomeration Schedule**

Stage	Cluster Combined		Coefficient	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	16	17	0.127	0	0	12
2	6	11	1.182	0	0	7
3	14	18	2.968	0	0	6
4	3	9	5.027	0	0	15
5	2	4	9.843	0	0	10
6	14	15	15.014	3	0	12
7	5	6	20.501	0	2	8
8	5	7	28.356	7	0	9
9	5	19	42.164	8	0	11
10	2	12	56.053	5	0	15
11	5	8	79.980	9	0	13
12	14	16	146.217	6	1	14
13	5	10	212.506	11	0	16
14	13	14	313.389	0	12	18
15	2	3	416.162	10	4	16
16	2	5	615.409	15	13	17
17	1	2	1102.758	0	16	18
18	1	13	2353.864	17	14	0

Selanjutnya untuk menentukan banyaknya *cluster* yang terbentuk dapat diketahui dengan cara melihat perubahan persentase pada jumlah kuadrat dalam *cluster*. Hasil perubahan persentase jumlah kuadrat dalam *cluster* dapat dilihat pada Tabel 3.

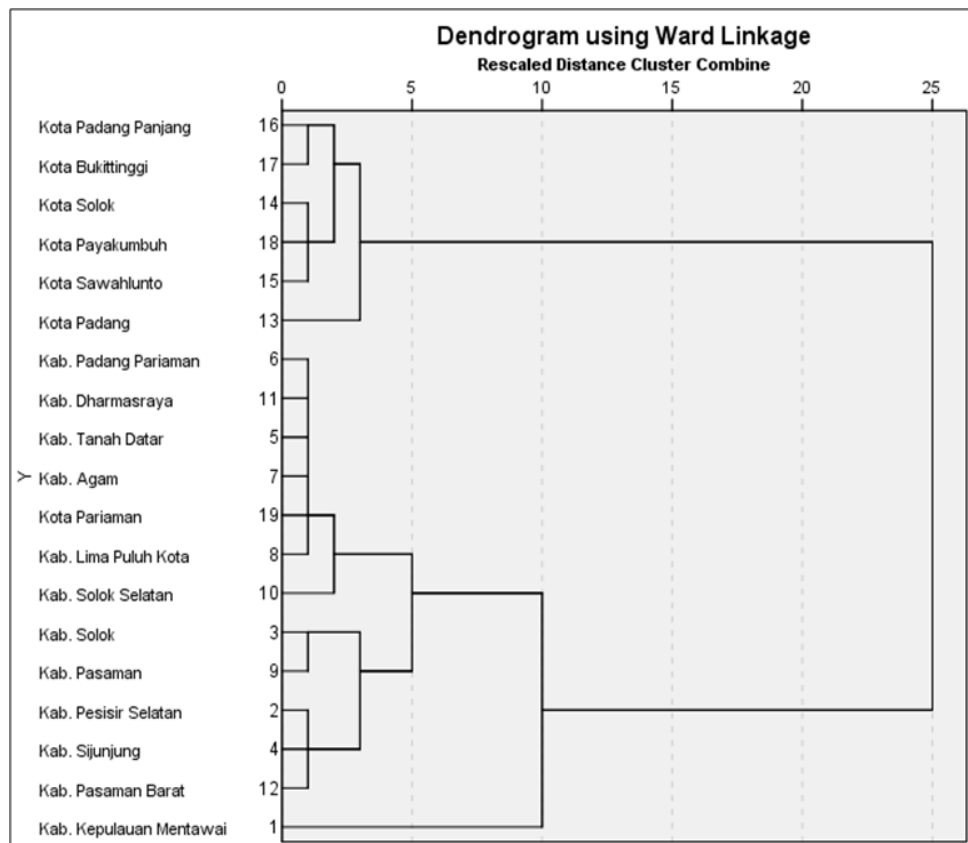
**Tabel 2. Hasil Perubahan Persentase Jarak Squared Euclidean**

Stage	Jumlah Kuadrat dalam Cluster	Perubahan Persentase
1	0.254	186%
2	2.110	146%
3	3.571	63%
4	4.199	543%
5	9.632	188%
6	7.756	48%

7	8.232	224%
8	10.472	679%
9	17.260	357%
10	20.834	788%
11	28.712	2649%
12	55.198	2214%
13	77.338	4372%
14	121.060	3542%
15	85.644	1733%
16	68.313	<b>45965%</b>
17	527.962	22321%
18	304.756	

Dari Tabel 3, dapat diketahui bahwa perubahan persentase yang tertinggi yaitu sebesar 45965% yang berada pada stage 16. Banyaknya *cluster* yang terbentuk yaitu hasil pengurangan dari banyaknya objek yang diamati ( $n$ ) dengan *stage* saat proses dihentikan. Banyaknya objek yang diamati adalah 19. Sehingga, banyaknya *cluster* yang terbentuk yaitu  $19-16 = 3$  *cluster*.

Adapun hasil *cluster* pengelompokan Kabupaten/Kota berdasarkan indikator ketenagakerjaan di Provinsi Sumatera Barat menggunakan metode *Ward* membentuk dendrogram yang dapat dilihat pada Gambar 1.



**Gambar 1. Hasil Cluster Menggunakan Metode Ward**

Dari Gambar 2, dapat diketahui bahwa hasil cluster yang diperoleh yaitu *cluster* 1 terdapat 1 anggota yang terdiri dari Kabupaten Kepulauan Mentawai. *Cluster* 2 terdapat 12 anggota yang terdiri dari Kabupaten Pesisir Selatan, Kabupaten Solok, Kabupaten Sijunjung, Kabupaten Tanah Datar, Kabupaten Padang Pariaman,

Kabupaten Agam, Kabupaten Lima Puluh Kota, Kabupaten Pasaman, Kabupaten Solok Selatan, Kabupaten Dharmasraya, Kabupaten Pasaman Barat dan Kota Pariaman. *Cluster* 3 terdapat 6 anggota yang terdiri dari Kota Padang, Kota Solok, Kota Sawahlunto, Kota Padang Panjang, Kota Bukittinggi dan Kota Payakumbuh.

Setelah diperoleh hasil *cluster* menggunakan metode *Ward*, maka untuk mengetahui karakteristik dari setiap *cluster* dapat dilakukan dengan menghitung rata-rata dari masing-masing variabel di setiap *clusternya*. Adapun nilai rata-rata dari masing-masing variabel di setiap *clusternya* dapat dilihat pada Tabel 4.

**Tabel 3. Interpretasi Hasil Cluster**

Variabel	Cluster 1	Cluster 2	Cluster 3	Total
TPT (X1)	1,39	5,28	5,92	5,28
TPAK (X2)	78,25	71,23	66,94	70,24
TKK (X3)	98,61	94,72	94,09	94,72
TPTP (X4)	58,26	37,08	21,99	33,43

Dari Tabel 4, dapat diketahui bahwa hasil cluster yang diperoleh yaitu *cluster* 1 memiliki jumlah rata-rata ketenagakerjaan diatas jumlah rata-rata ketenagakerjaan di Provinsi Sumatera Barat. Hal ini menunjukkan bahwa Cluster 1 merupakan *cluster* dengan tingkat ketenagakerjaan yang tinggi. *Cluster* 2 memiliki jumlah rata-rata ketenagakerjaan yang hampir sama dengan jumlah rata-rata ketenagakerjaan di Provinsi Sumatera Barat. Hal ini menunjukkan bahwa *cluster* 2 merupakan cluster dengan tingkat ketenagakerjaan yang sedang. *Cluster* 3 memiliki jumlah rata-rata ketenagakerjaan dibawah jumlah rata-rata ketenagakerjaan di Provinsi Sumatera Barat. Hal ini menunjukkan bahwa *cluster* 3 merupakan cluster dengan tingkat ketenagakerjaan yang rendah. Untuk mengetahui seberapa baik hasil *cluster* yang diperoleh dapat dilakukan validasi *cluster*. Pada penelitian ini validasi yang digunakan yaitu *Dunn Index*. Adapun hasil nilai *Dunn Index* sebagai berikut:

**Tabel 4. Nilai Dunn Index**

Jumlah Cluster	Dunn Index
2	0.359
<b>3</b>	<b>0.780</b>
4	0.295
5	0.219

Dari **Tabel 5**, dapat diketahui bahwa nilai *Dunn Index* yang tertinggi yaitu pada jumlah *cluster* 3 dengan nilai *Dunn Index* sebesar 0.780. Artinya, jumlah cluster yang terbaik untuk hasil pengelompokan menggunakan metode *Ward* sebanyak 3 cluster.

## SIMPULAN

Berdasarkan hasil dan pembahasan yang telah dipaparkan pada penelitian ini, maka dapat disimpulkan bahwa dalam pengelompokan Kabupaten/Kota di Provinsi Sumatera Barat Tahun 2022 berdasarkan indikator ketenagakerjaan menggunakan metode *Ward* terbentuk menjadi 3 *Cluster* dengan karakteristik sebagai berikut: *Cluster* 1 terdapat 1 anggota yang terdiri dari Kabupaten Kepulauan Mentawai yang termasuk ke dalam *cluster* dengan tingkat ketenagakerjaan tinggi. *Cluster* 2 terdapat 12 anggota yang terdiri dari Kabupaten Pesisir Selatan, Kabupaten Solok, Kabupaten Sijunjung, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten Lima Puluh Kota, Kabupaten Pasaman, Kabupaten Solok Selatan, Kabupaten Dharmasraya, Kabupaten Pasaman Barat dan Kota Pariaman yang termasuk ke dalam *cluster* dengan tingkat ketenagakerjaan sedang. *Cluster* 3 terdapat 6 anggota yang terdiri dari Kota Padang, Kota Solok, Kota Sawahlunto, Kota Padang Panjang, Kota Bukittinggi dan Kota Payakumbuh yang termasuk ke dalam *cluster* dengan tingkat ketenagakerjaan rendah.

## DAFTAR PUSTAKA

- Badan Pusat Statistik (BPS) Provinsi Sumatera Barat. 2023. *Provinsi Sumatera Barat dalam Angka 2023*. Padang : BPS Provinsi Sumatera Barat.
- Brock, Guy, dkk. 2008. cIValid: An R Package for cluster Validation. *Ournal of Statistical Software*. 25(4): 1-22.
- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. 2010. *Multivariate Data Analysis*. 7<sup>th</sup> Edition, New York: Perason.
- Han, J., Kamber, M., & Pei, J. 2012. *Data Mining : Concept and Techniques*, 3<sup>d</sup> edition. Amsterdam: Morgan Kaufmann-Elsevier.
- Johnson, R.A & Wichern, D.W. 2002. *Applied Multivariate Statistical Analysis Fifth Edition*. New Jersey : Prentice Hall International.
- \_\_\_\_\_. 2007. *Applied Multivariate Statistical Analysis Sixth Edition*. United States: Prentice Hall International.
- Nugroho, S. 2008. *Statistika Multivariat Terapan, Edisi Pertama*. Bengkulu : UNIB Press.
- Dewi, L.S, Talakua, M.W., Lesnuss, Y.A., & Matdoan, M.Y. (2021). Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Pendidikan dengan Menggunakan Metode Ward. *Jurnal Statistika dan Aplikasinya*, 5(1), 51 – 60.
- Rahmawati, D.N., & Wijayanto, A.W. (2023). Perbandingan Algoritma Partitioning dan Hierarchical Clustering untuk Pengelompokan Wilayah Menurut Karakteristik Pengangguran di Pulau Jawa Tahun 2021. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 10(3), 265 – 279.
- Govender, P., & Sivakumar, V. 2019. Application of K-Means and Hierarchical Clustering Techniques for Analysis of Air Pollution : A Review (1980-2019). *In Atmospheric Pollution Research*.
- Imasdiani., Purnamasari, I., & Amijaya, F.D.T. (2022). Perbandingan Hasil Analisis Cluster Dengan Menggunakan Metode Average Linkage Dan Metode Ward (Studi Kasus : Kemiskinan Di Provinsi Kalimantan Timur Tahun 2018). *Jurnal Eksponensial*, 13(1), 9 – 18.