**Journal of Data Analysis**
Vol.7, No.2, December 2024, p. 81-87

# *Clustering Analysis* Terhadap Kondisi Tingkat Pernikahan dan Perceraian Indonesia

**Rizka Pitri[1*], Linda Rassiyanti[2]**

[1] Program Studi Sains Data, Fakultas Sains dan Teknologi, UIN Raden Intan, Lampung, Indonesia
[2] Program Studi Sains Data,Fakultas Sains, Institut Teknologi Sumatera, Lampung, Indonesia
E-mail: rizka@radenintan.ac.id*, linda.rassiyanti@sd.itera.ac.id

**Abstrak**

Menikah dan hidup bahagia adalah impian hampir semua orang. Selama menjalani rumah tangga perselisihan pasti terjadi namun semua berharap perselisihan tidak berakhir pada perceraian. Data terbaru dari Badan Pusat Statistik (BPS) menunjukkan bahwa angka pernikahan dan perceraian bervariasi di setiap provinsi, mencerminkan dinamika sosial yang berbeda-beda di setiap daerah. *Clustering* untuk data pernikahan dan perceraian di Indonesia dapat menggunakan metode *single linkage* dan *k-means.* Berdasarkan hasil yang diperoleh, menunjukan bahwa jumlah pernikahan di berbagai provinsi di Indonesia menunjukkan adanya variasi yang cukup signifikan dan terdapat tiga *cluster* yang dapat di-*clustering* berdasarkan jumlah pernikahan dan perceraian.

**Informasi Artikel**

***Sejarah Artikel:***
Diajukan 30 Desember 2024
Diterima 2 Januari 2025

***Kata Kunci :***
*K-Means*
*Perceraian*
*Pernikahan*
*Single Linkage*

**Abstract**

Marriage and living a happy life are a dream shared by almost everyone. In married life, disputes are inevitable, but everyone hopes that these disagreements will not lead to divorce. The latest data from the Central Bureau of Statistics (BPS) show that marriage and divorce rates vary across provinces, reflecting the diverse social dynamics in different regions. The grouping of marriage and divorce data in Indonesia can be conducted using the single linkage and k-means methods. Based on the results obtained, it is evident that the number of marriages across various provinces in Indonesia shows varies significantly, and three distinct groups can be identified based on the number of marriages and divorces.

***Keywords:***
*K-Means*
*Divorce*
*Married*
*Single Linkage*

## 1. Pendahuluan

Menikah dan hidup bahagia adalah impian hampir semua orang. Selama menjalani rumah tangga perselisihan pasti terjadi namun semua berharap perselisihan tidak berakhir pada perceraian. Data terbaru dari Badan Pusat Statistik (BPS) menunjukkan bahwa angka pernikahan dan perceraian bervariasi di setiap provinsi, mencerminkan dinamika sosial yang berbeda-beda di setiap daerah. Pada tahun 2023, misalnya, Sumatera Utara mencatat 70.630 pernikahan dengan 18.269 perceraian, sedangkan Sumatera Barat memiliki 38.355 pernikahan dan 9.266 perceraian [1].

Metode *Single Linkage* merupakan metode hierarki yang paling sederhana, dimana grup dibentuk dari entitas individu dengan menggabungkan individu terdekat, di mana istilah individu terdekat berarti jarak terkecil atau *similarities* terbesar. *Complete Linkage* berlangsung dengan cara yang hampir sama seperti *Single Linkage*, dengan satu pengecualian penting, yaitu pada setiap tahap, jarak (*similarities*) antar klaster ditentukan oleh jarak (*similarities*) antara dua elemen, satu dari setiap klaster, yaitu yang paling jauh.

Dengan demikian, *Complete Linkage* memastikan bahwa semua amatan dalam sebuah klaster berada dalam jarak maksimum (atau *similarities* minimum) satu sama lain. Sedangkan, *Average Linkage* memperlakukan jarak antara dua klaster sebagai jarak *Euclidean* rata-rata antara semua pasangan item di mana salah satu anggota dari pasangan dimiliki oleh setiap klaster [2].

Metode *K-means* merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, *K-means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk dari metode *K-means* ini sangatlah tergantung pada inisiasi nilai pusat awal *cluster* yang diberikan. Hal ini menyebabkan hasil *cluster* yang terbentuk merupakan solusi yang sifatnya lokal optimal. Oleh karena itu, *K- means* dikolaborasikan oleh metode hierarki untuk penentuan pusat awal *cluster* [3].

## 2. Tinjauan Kepustakaan

### *2.1. Analisis Cluster*

Analisis *cluster* merupakan suatu metode yang membagi individu kedalam kelompok yang bermakna. Analisis *cluster* mengelompokkan obyek berdasarkan informasi yang diperoleh pada data yang menggambarkan obyek dan keterkaitannya [4]. Analisis *cluster* adalah model multivariat yang dibuat dengan tujuan mengklasifikasikan data menurut karakteristik data tersebut.

Analisis *cluster* memiliki dua jenis *clustering*, yaitu metode hierarki dan non-hierarki. Metode hierarki adalah metode yang dirancang untuk dekomposisi berhierarki (tingkatan) dari kumpulan data menurut karakteristik pada data tersebut. Hal tersebut bertolakbelakang dengan metode non-hierarki, yaitu metode yang digunakan untuk mengelompokkan data dan jumlah cluster yang akan dibuat dapat ditetapkan sebelumnya. Ada berbagai jenis analisis *cluster* yang digunakan dalam metode hierarki, yaitu *single linkage* dan complete *linkage*. Banyak metode *clustering* non-hierarki yang sering digunakan, seperti metode *k-means* dan *k-medoids* [5].

Metode *single linkage* merupakan bagian dari metode *hierarchical clustering* dimana pengelompokkannya dengan mencari jarak paling dekat atau minimum diantara pengamatan suatu objek. Rumus yang digunakan untuk menentukan jarak antara dengan klaster lainnya adalah seperti berikut [6].

$$d_{(UV)k} = min\,\{d_{uk}, d_{vk}\} \tag{1}$$

keterangan:
$d_{uk}$ : jarak antara objek $u$ dengan objek $k$
$d_{vk}$ : jarak antara objek $v$ dengan objek $k$
$d_{(UV)k}$ : jarak paling minimum antara objek dengan $k$

Analisis *cluster* harus memenuhi dua asumsi yaitu:
1. Sampel yang diambil harus benar-benar dapat mewakili populasi.
2. Tidak ada multikolinieritas (korelasi antar objek), jika ada maka nilai multikolinieritas kurang dari 0,5 [4].

Metode *single linkage* merupakan bagian dari metode *hierarchical clustering* dimana pengelompokkannya dengan mencari jarak paling dekat atau minimum diantara pengamatan suatu objek. *K-Means clustering* merupakan bagian dari metode *hierarchical clustering* yang termasuk pada pendekatan partisi. Tujuannya adalah untuk mencari jumlah jarak yang paling kecil diantara suatu titik data. Penentuan jumlah *cluster* pada metode ini harus di inisialisasikan terlebih dahulu sebelum dilakukan pengelompokkan [6].

## 3. Metodologi Penelitian

Penelitian ini menggunakan data nikah dan cerai menurut provinsi di Indonesia pada tahun 2023, yang terdiri dari 34 provinsi. Data yang diperoleh bersumber dari data Kependudukan dan Migrasi Badan Pusat Statistik (BPS). Terdapat 4 peubah dari data tersebut yang terdiri dari:

**Tabel 1** Peubah

| No | Nama Peubah | Jenis Peubah | Keterangan | Satuan |
|---|---|---|---|---|
| 1 | Nikah | Peubah Bebas ($X_1$) | Jumlah pernikahan yang tercatat di KUA | Kasus |
| 2 | Cerai Talak | Peubah Bebas ($X_2$) | Jumlah perceraian yang diajukan suami | Kasus |
| 3 | Cerai Gugat | Peubah Bebas ($X_3$) | Jumlah perceraian yang diajukan istri | Kasus |
| 4 | Jumlah Cerai | Peubah Terikat $X_4$) | Total perceraian (cerai talak+ceria gugat) | Kasus |

Penelitian ini menggunakan dua metode c*lustering*, yaitu *single linkage* dan *k-means clustering*. Metode *single lingkage* digunakan untuk menentukan jumlah *clustering* awal berdasarkan provinsi. *Clustering* ini dibentuk berdasarkan kemiripan pola pernikahan dan perceraian. Setelah jumlah *clustering* ditentukan melalui metode *single linkage* maka jumlah *cluster* tersebut digunakan pada metode *K-means clustering*.

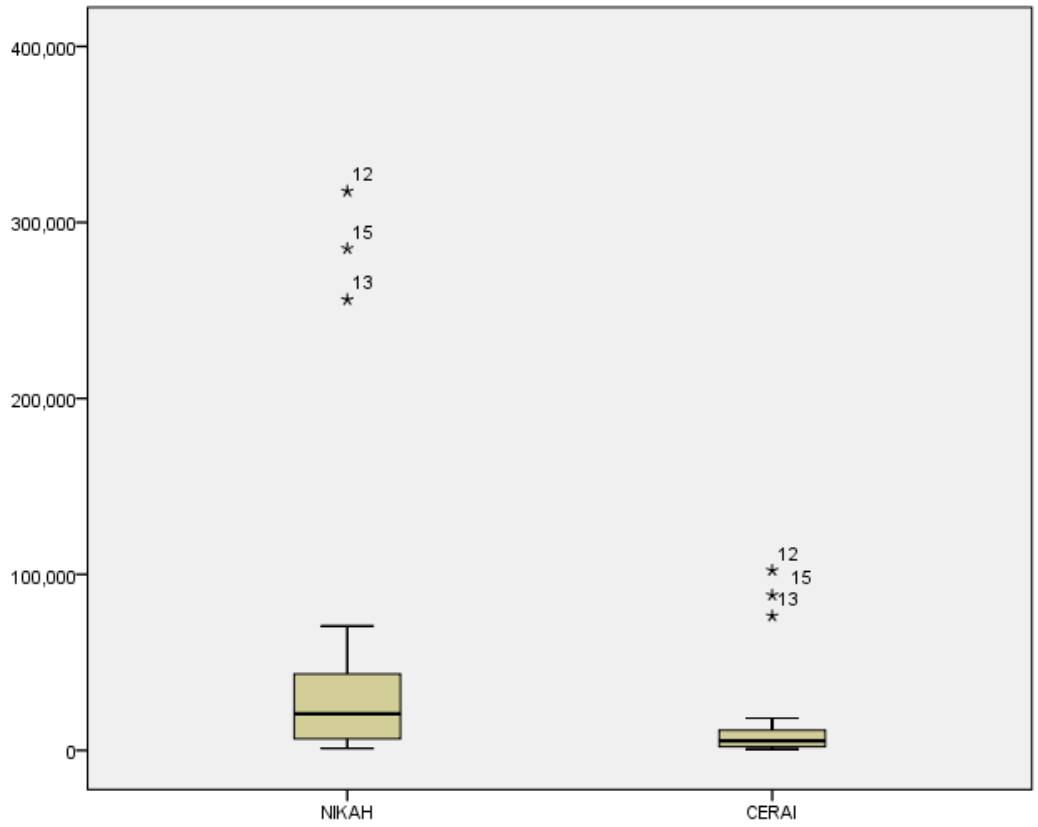
## 4. Analisis Pembahasan

Tahun 2023, sebanyak 1.577.255 pasangan di Indonesia yang melakukan pernikahan. Provinsi Jawa Barat merupakan provinsi yang memiliki jumlah pernikahan tertinggi selama tahun 2023, yaitu sebanyak 317.715 pasangan. Tak hanya jumlah pernikahan tertinggi yang terjadi di Jawa Barat, namun angka perceraian pun tertinggi terjadi di Jawa Barat di tahun 2023, yaitu sebanyak 102.280 pasangan. Sedangkan Provinsi Papua Selatan memiliki angka pernikahan terendah, yaitu 871 pasangan. Walaupun terdapat 1.577.255 pasangan yang menikah di tahun 2023, namun terdapat 29.4% pasangan yang melakukan perceraian, atau setara 463.654 pasangan. Mayoritas dari pasangan bercerai tersebut melakukan perceraian jenis gugatan, yaitu sebanyak 353.403. Sisanya sebanyak 111.251 pasangan melakukan perceraian dikarenakan talak yang dilakukan oleh suami. Nusa Tenggara Timur memiliki jumlah pasangan bercerai terendah dibandingkan provinsi lainnya, yaitu sebanyak 621 pasangan atau setara 21.2% dari jumlah menikah di tahun 2023.

**Tabel 2** Ringkasan data nikah dan cerai tahun 2023

| Peubah | Minimum | Maksimum | Rataan | Simpangan Baku |
|---|---|---|---|---|
| Nikah | 871 | 317.715 | 42.628,51 | 76.273,910 |
| Cerai | 621 | 102.280 | 13.636,88 | 24.538,804 |

Jika dilihat dari persebaran data nikah tahun 2023 pada Gambar 1, 25% provinsi memiliki angka nikah diatas 41.832 sedangkan angka perceraian diatas 5.428. Jika dilihat berdasarkan provinsinya maka terdapat Sembilan provinsi yang memiliki angka pernikahan di atas 41.832, yaitu Provinsi DKI Jakarta, Sumatera Selatan, Lampung, Banten, Sumatera Utara, Jawa Tengah, Jawa Timur, dan Jawa Barat. Akan tetapi, jika dilihat dari angka perceraian maka terdapat 17 provinsi yang memiliki angka perceraian diatas 5.428, yaitu Kalimantan Barat, Yogyakarta, Aceh, Kalimantan Selatan, Kalimantan Timur, NTB, Sumatera Barat, Riau, Sumatera Selatan, Sulawesi Selatan, Lampung, Banten, DKI Jakarta, Sumatera Utara, Jawa Timur, Jawa Tengah, dan Jawa Barat.

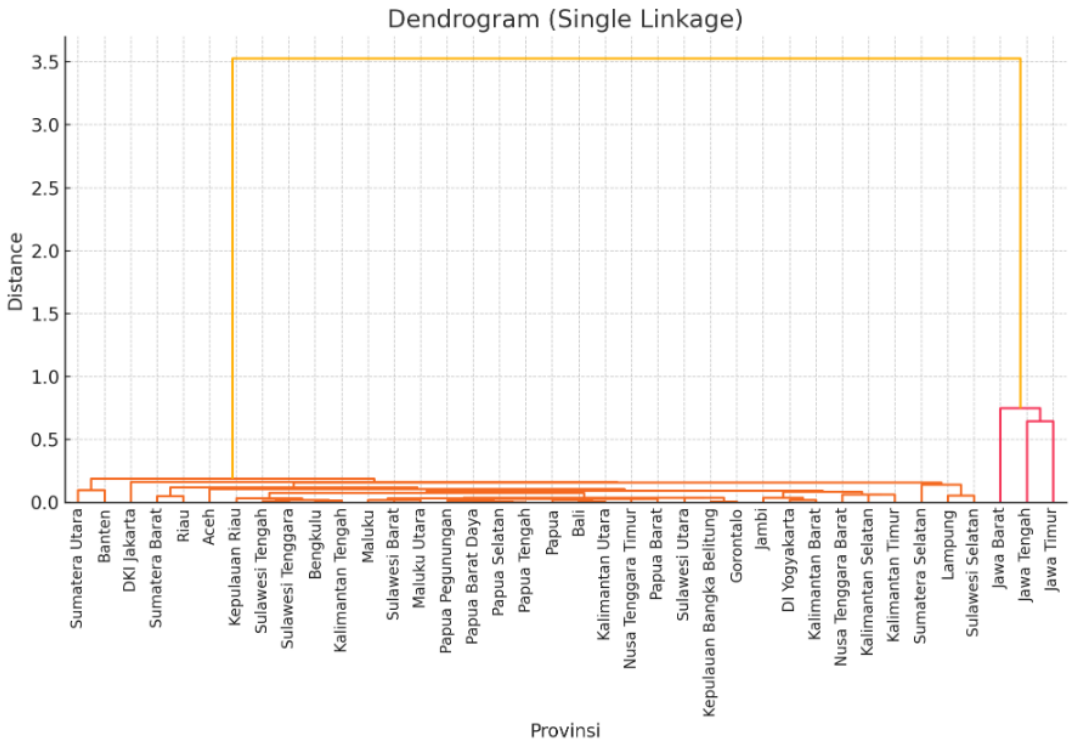



**Gambar 1** Sebaran angka nikah dan cerai tahun 2023.

*Clustering* pada provinsi berdasarkan kemiripan jumlah pernikahan dan perceraian yang ditunjukkan pada Gambar 2. Beberapa provinsi seperti Aceh, Bengkulu, dan Kepulauan Bangka Belitung memiliki pola yang sangat mirip, yakni angka pernikahan dan perceraian yang relatif rendah. Di sisi lain, provinsi-provinsi seperti Jawa Barat, Jawa

Timur, dan Jawa Tengah memiliki karakteristik yang cukup berbeda, dengan angka pernikahan dan perceraian yang cenderung lebih tinggi dibandingkan provinsi lain.

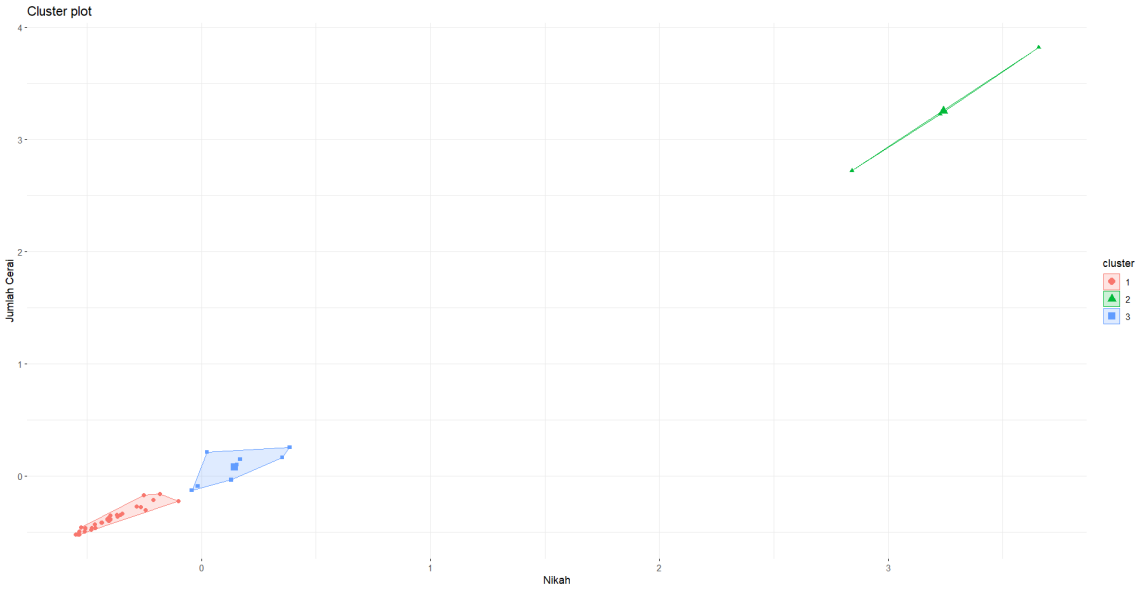
Jika dendrogram dipotong pada jarak tertentu, terlihat bahwa provinsi-provinsi ini dapat dikelompokkan menjadi beberapa *cluster* utama. *Cluster* pertama terdiri dari provinsi-provinsi dengan angka pernikahan dan perceraian yang rendah, sementara *cluster* kedua mencakup provinsi-provinsi dengan angka sedang hingga tinggi. Beberapa provinsi seperti DKI Jakarta dan Sumatera Utara menunjukkan posisi penggabungan yang berada di antara dua kelompok tersebut, menandakan bahwa karakteristiknya tidak sepenuhnya mirip dengan kelompok manapun.




**Gambar 2** Dendogram *Single Linkage.*

Provinsi dikelompokkan ke dalam tiga *cluster* berdasarkan jumlah pernikahan (nikah) dan jumlah perceraian (jumlah cerai) yang ditunjukkan pada Gambar 3. *Cluster* satu (merah), mencakup provinsi-provinsi dengan jumlah pernikahan dan perceraian yang relatif rendah dibandingkan dengan *cluster* lainnya. Provinsi-provinsi dalam *cluster* ini cenderung memiliki karakteristik demografis yang lebih seragam, dengan tingkat aktivitas pernikahan dan perceraian yang kecil. *Cluster* dua (hijau) merupakan kelompok dengan jumlah pernikahan dan perceraian yang paling tinggi. Provinsi dalam *cluster* ini memiliki angka pernikahan yang sangat besar, diikuti dengan angka perceraian yang signifikan. Hal ini mungkin terkait dengan populasi yang lebih besar atau aktivitas sosial yang lebih intens di wilayah tersebut. *Cluster* 3 (biru) merupakan kelompok dengan jumlah pernikahan dan perceraian yang sedang, tidak terlalu rendah seperti *cluster* satu, tetapi juga tidak sebesar *cluster* dua. Provinsi dalam *cluster* ini menunjukkan pola aktivitas sosial yang moderat, yaitu tidak memprioritaskan pernikahan dan lebih

berfokuskan kepada aktivitas social lainnya seperti melakukan permberdayaan masyarakat di lingkungan rumah maupun perkerjaan.




**Gambar 3** Hasil Clustering *K-Means.*

Tabel 3 merupakan klasifikasi provinsi berdasarkan *cluster* yang diperoleh. Jawa Barat, Jawa Tengah, dan Jawa Timur merupakan provinsi yang memiliki jumlah pernikahan dan perceraian paling tinggi dibandingan provinsi lainnya. Hal ini bisa disebabkan jumlah populasi di provinsi tersebut lebih tinggi dibandingkan provinsi lainnya.

**Tabel 3** Klasifikasi provinsi berdasarkan *cluster*

| No | Provinsi | *Cluster* | No | Provinsi | *Cluster* |
|---|---|---|---|---|---|
| 1 | Aceh | 1 | 20 | Maluku | 1 |
| 2 | Jambi | 1 | 21 | Maluku Utara | 1 |
| 3 | Bengkulu | 1 | 22 | Papua Barat | 1 |
| 4 | Kepulauan Bangka Belitung | 1 | 23 | Papua Barat Daya | 1 |
| 5 | Kepulauan Riau | 1 | 24 | Papua | 1 |
| 6 | DI Yogyakarta | 1 | 25 | Papua Selatan | 1 |
| 7 | Bali | 1 | 26 | Papua Tengah | 1 |
| 8 | Nusa Tenggara Barat | 1 | 27 | Papua Pegunungan | 1 |
| 9 | Nusa Tenggara Timur | 1 | 28 | Jawa Barat | 2 |
| 10 | Kalimantan Barat | 1 | 29 | Jawa Tengah | 2 |
| 11 | Kalimantan Tengah | 1 | 30 | Jawa Timur | 2 |
| 12 | Kalimantan Selatan | 1 | 31 | Sumatera Utara | 3 |
| 13 | Kalimantan Timur | 1 | 32 | Sumatera Barat | 3 |
| 14 | Kalimantan Utara | 1 | 33 | Riau | 3 |
| 15 | Sulawesi Utara | 1 | 34 | Sumatera Selatan | 3 |
| 16 | Sulawesi Tengah | 1 | 35 | Lampung | 3 |
| 17 | Sulawesi Tenggara | 1 | 36 | DKI Jakarta | 3 |

| No | Provinsi | *Cluster* | No | Provinsi | *Cluster* |
|---|---|---|---|---|---|
| 18 | Gorontalo | 1 | 37 | Banten | 3 |
| 19 | Sulawesi Barat | 1 | 38 | Sulawesi Selatan | 3 |

## 5. Kesimpulan

Berdasarkan penelitian yang telah dilakukan diperoleh beberapa kesimpulan, yaitu:

1. Data penikahan dan perceraian di Indonesia memiliki tiga kelompok.
2. Kelompok provinsi dengan jumlah pernikahan dan perceraian terendah masuk dalam *cluster* satu.
3. Kelompok provinsi dengan jumlah pernikahan dan perceraian terbanyak masuk dalam *cluster* dua.
4. Kelompok provinsi dengan jumlah pernikahan dan perceraian sedang masuk dalam *cluster* tiga.

## Daftar Pustaka

[1] Purwaningsih, E., & Nurelasari, E. 2023. Implementasi Metode K-Means Clustering Dengan Davies Bouldin Index Pada Analisis Faktor Penyebab Perceraian. 7(2):134–143.

[2] Syafiyah, U., Asrafi, I., Wicaksono, B., Puspitasari, D. P., & Sirait, F. M. 2020. Analisis Perbandingan Hierarchical dan Non-Hierarchical Clustering Pada Data Indikator Ketenagakerjaan di Jawa Barat Tahun 2020. 2020.803–812.

[3] Alfina, T., Santosa, B., & Barakbah, R. 2012. *Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus: Problem Kerja Praktek Jurusan Teknik Industri ITS)*.

[4] Wahyuni, R., Nugroho, S., Novianti, P., Statistika, M. J., Matematika, P., Pembimbing, D., Statistika, J., & Matematika, P. 2013. Analisis klaster dengan Menggunakan Metode Single Linkage dan metode K-means (Studi Kasus: Data Produksi, Produktivitas, dan Luas Panen Tanaman Sayuran Di Provinsi Bengkulu Tahun 2013. 1–9.

[5] Alamtaha, Z., Djakaria, I., & Yahya, N. I. 2023. Implementasi Algoritma Hierarchical Clustering dan Non-Hierarchical Clustering untuk Pengelompokkan Pengguna Media Sosial. 4(1):33–43. https://doi.org/10.20956/ejsa.vi.24830.

[6] Parulian, R., & Butar, B. 2023. Analisis Hierarchical Dan Non-Hierarchical C lustering Untuk Pengelompokkan Potensi Ekonomi Kelautan Indonesia 2021 Hierarchical and Non-Hierarchical Clustering Analysis for Classifying Indonesia's Sea Economic Potential in 2021. 11(3):543–553. https://doi.org/10.26418/justin.v11i3.67283.