Pengelompokan Perangkat Daerah di Jawa Barat Berdasarkan Jenis Pajak Menggunakan Metode *Ward*

Dyah Listya Rini*, Abdul Kudus, Suwanda

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*rdyahlistya@gmail.com, abdul.kudus@unisba.ac.id, suwanda@unisba.ac.id

Abstract. Cluster analysis is a multivariate statistical analysis to classify several objects with similar characteristics. Hierarchical cluster analysis is a method that is carried out by creating a level based on the same characteristics. Hierarchical cluster analysis with the Ward method is a method that aims to minimize the increase in variation between objects in one cluster. On this occasion, the author obtained secondary data from BPKAD in the form of recapitulated tax collection data in SIPD in January-August 2023. Based on the grouping carried out, 2 clusters were obtained with the number of cluster 1 members as many as 1 regional apparatus and the number of cluster 2 members as many as 37 regional apparatus. The final result of the grouping is that the Education Office is the only member in cluster 1 and the Health Office and 36 other regional apparatus are in cluster 2. The average value of each type of tax provided by cluster 1 is greater than the average value of each type of tax provided by cluster 2, so it can be said that cluster 1 has the greatest tax attraction.

Keywords: *Clustering, Ward, Tax.*

Abstrak. Analisis klaster merupakan analisis statistika multivariat untuk mengklasifikasikan beberapa objek dengan karakteristik yang sama. Analisis klaster hierarki merupakan metode yang dilakukan dengan membuat suatu tingkatan berdasar karakteristik yang sama. Analisis klaster hierarki dengan metode Ward merupakan metode yang bertujuan meminimalkan peningkatan variasi antar objek yang ada dalam satu klaster. Pada kesempatan kali ini penulis memperoleh data sekunder dari pihak instansi terkait berupa data rekapitulasi tarikan pajak di SIPD pada bulan Januari-Agustus tahun 2023. Berdasarkan pengelompokan yang dilakukan diperoleh sebanyak 2 klaster dengan jumlah anggota klaster 1 sebanyak 1 perangkat daerah dan jumlah anggota klaster 2 sebanyak 37 perangkat daerah. Hasil akhir pengelompokannya yakni Perangkat Daerah 1 ($\boldsymbol{PD_1}$) sebagai satu-satunya anggota dalam klaster 1 dan Perangkat Daerah 2 ($\boldsymbol{PD_2}$) serta 36 perangkat daerah lainnya berada dalam klaster 2. Nilai rata-rata tiap jenis pajak yang diberikan oleh klaster 1 lebih besar dibandingkan nilai rata-rata tiap jenis pajak yang diberikan oleh klaster 2, sehingga dapat dikatakan bahwa klaster 1 memiliki tarikan pajak paling besar.

Kata Kunci: *Pengelompokan, Ward, Pajak.*

Corresponding Author
Email: abdul.kudus@unisba.ac.id

A. Pendahuluan

Analisis klaster merupakan suatu metode yang tepat dalam mengidentifikasi objek-objek yang homogen membentuk beberapa kelompok yang disebut klaster [1]. Karena tujuan analisis ini merupakan mendapatkan klaster yang sehomogen mungkin, maka dapat digunakan kesamaan skor nilai yang dianalisis sebagai dasar dalam melakukan pengklasteran. Data mengenai ukuran kesamaan tersebut dapat dikelompokkan dengan analisis klaster sehingga dapat ditentukan setiap anggota klasternya. Secara garis besar metode dalam melakukan analisis klaster dibagi menjadi 2 yaitu metode hierarki dan non hierarki.

Analisis klaster dengan metode non hierarki atau dapat disebut dengan pengklasteran berbasis partisi merupakan suatu metode yang digunakan untuk menghasilkan partisi dari sebuah data [1]. Hal tersebut akan membuat suatu objek atau karakteristik dalam 1 klaster menjadi lebih mirip satu dengan yang lainnya [2]. Analisis klaster dengan metode non hierarki dilakukan dengan menentukan jumlah klasternya terlebih dahulu, dan tidak dilakukan secara bertahap. Sedangkan metode hierarki umumnya digunakan pada ukuran sampel yang kecil dan jumlah kelompok yang hendak dibentuk belum diketahui. Pengelompokan disajikan dalam bentuk dendrogram yang mirip dengan diagram pohon. Metode hierarki dibagi menjadi 2 yaitu metode penggabungan (agglomerative) atau metode pembagian (divisive). Pada metode aglomerasi atau penggabungan, setiap penelitian awalnya diasumsikan sebagai suatu klaster tersendiri dimana terdapat klaster yang sesuai dengan jumlah penelitian. Kemudian, dua klaster yang memiliki kesamaan terdekat digabungkan menjadi klaster baru sehingga jumlah klaster terus berkurang satu di setiap tahapannya. Sebaliknya, apabila digunakan metode pemecahan dimulai dari satu klaster besar yang mencakup semua observasi. Selain itu, observasi yang berbeda dipisahkan dan dikelompokkan menjadi suatu kelompok-kelompok kecil. Proses ini akan terus berlanjut hingga setiap observasi membentuk kelompok masing-masing [1]. Terdapat 5 metode Agglomerative pada analisis klaster hierarki, salah satunya adalah metode Ward. Analisis klaster hierarki dengan metode Ward merupakan metode yang bertujuan meminimalkan peningkatan variasi antar objek yang ada dalam satu klaster [3].

Berdasarkan Undang-Undang Nomor 42 Tahun 2009 tentang Perubahan Ketiga Atas Undang-Undang Nomor 8 Tahun 1983 tentang Pajak Pertambahan Nilai Barang dan Jasa serta Pajak Penjualan Atas Barang Mewah, PPN merupakan pajak atas konsumsi barang dan jasa di Daerah Pabean yang dikenakan secara bertingkat di setiap tahapan rantai produksi dan distribusi. Sementara itu, Rahayu dan Suhayati (2013) dalam Mahardika (2022) juga menyebutkan bahwa pajak pertambahan nilai adalah pajak atas setiap pertambahan nilai yang muncul karena terdapat elemen produksi yang terpakai pada proses produksi hingga distribusi barang kepada konsumen [4].Pajak Penghasilan Pasal 21 atau dikenal dengan sebutan PPh 21 adalah pajak atas penghasilan berupa gaji, upah, honorarium, tunjangan, dan biaya lainnya yang dikenakan kepada sesuatu yang berkaitan dengan pekerjaan atau jabatan, jasa dan kegiatan yang dilakukan oleh pengusaha kena pajak dalam negeri. (Kementerian Keuangan Republik Indonesia [5].

Berdasarkan Undang-Undang Nomor 36 Tahun 2008 tentang Pajak Penghasilan, PPh 22 adalah pajak yang berkaitan dengan kegiatan perdagangan barang seperti ekspor, impor, dan re-impor maupun kegiatan perdagangan lain yang pemungutannya dilakukan oleh bendaharawan atau badan usaha tertentu terhadap wajib pajak [6]. Pajak Penghasilan Pasal 23 atau PPh 23 adalah pajak penghasilan atas modal, penyerahan jasa, bunga, dividen, royalti, atau hadiah dan penghargaan, selain yang telah dipotong pada PPh Pasal 21. Pembayaran PPh 23 dilakukan oleh pihak pemotong melalui bank persepsi yang telah ditunjuk oleh Kementerian Keuangan dengan jatuh tempo pada tanggal 10 bulan berikutnya [7]. PPh Pasal 4 ayat (2) adalah pajak atas beberapa jenis penghasilan yang dibayarkan sehubungan dengan jasa dan sumber tertentu yang tidak dapat dikreditkan dengan pajak penghasilan terutang dan dikenakan pada wajib pajak badan maupun wajib pajak orang pribadi. PPh Pasal 4 ayat (2) dikenal juga dengan sebutan PPh Final, maksudnya ialah pemotongan pajak hanya dilakukan sekali dalam sebuah masa pajak atau pada saat transaksi dengan pertimbangan kemudahan, kesederhanaan, kepastian, pengenaan pajak yang tepat waktu, dan pertimbangan lainnya [8].

Berdasarkan data yang ada, penulis ingin melakukan pengelompokan untuk seluruh

perangkat daerah yang ada di Provinsi Jawa Barat menjadi beberapa klaster. Hal ini dilakukan agar identifikasi karakteristik tiap perangkat daerah lebih jelas, rinci, dan mudah dibandingkan berdasarkan tarikan pajak tertinggi dan terendah. Pengelompokan ini dilakukan dengan menggunakan metode Ward.

Berdasarkan latar belakang yang telah diuraikan, maka tujuan dari penelitian adalah mengetahui pengelompokan perangkat daerah berdasarkan jumlah total tarikan pajak di SIPD bulan Januari-Agustus 2023 menggunakan analisis klaster hierarki metode Ward. Kemudian, diharapkan dari penelitian ini diharapkan dapat memberikan berbagai manfaat bagi penulis maupun pembaca sebagai informasi tambahan mengenai jenis perpajakan dan menambah wawasan mengenai analisis klaster terutama untuk metode hierarki Ward.

B. Metodologi Penelitian

Peneliti menggunakan metode teknik analisis klaster dengan menggunakan pendekatan kuantitatif. Populasi yang dipilih dalam penelitian ini adalah perangkat daerah di Jawa Barat dengan jumlah keseluruhan sebanyak 38 perangkat daerah. Data penelitian yang diperoleh merupakan data sekunder. Data yang digunakan adalah data tarikan pajak di Sistem Informasi Pemerintah Daerah (SIPD) pada perangkat daerah di Jawa Barat bulan Januari sampai Agustus 2023 yang diperoleh dari instansi terkait. Unit pengamatan dari penelitian ini adalah perangkat daerah yang terdiri dari 38 perangkat daerah. Jenis pajak merupakan variabel yang digunakan pada penelitian diantaranya terdiri dari Pajak Pertambahan Nilai (PPN), Pajak Penghasilan Pasal 21 (PPh 21), Pajak Penghasilan Pasal 22 (PPh 22), Pajak Penghasilan Pasal 23 (PPh 23), dan Pajak Penghasilan Pasal 4 ayat (2).

Adapun tahapan dari metode *Ward* yaitu:

1. Menghitung jarak antar objek (perangkat daerah) dengan menggunakan rumus jarak *Euclidean* seperti pada persaman berikut:

$$d_{ij} = \sqrt{\Sigma_{k=1}^{p}\left(x_{ik} - x_{jk}\right)^2},\ k = 1,2,\dots,p. \quad \dots (1)$$

   Keterangan:
   x_{ik} : Objek ke *i* pada variabel ke *k*.
   x_{jk} : Objek ke *j* pada variabel ke *k*.
   d_{ij} : Jarak *Euclidean* objek ke *i* dan objek ke *j*.

2. Membentuk klaster pertama dengan cra menentukan dua objek dengan nilai SSE terkecil menggunakan persamaan berikut:

$$SSE_{ij} = \frac{1}{2}d^2(x_i, x_j) \quad \dots (2)$$

3. Menentukan jarak terdekat antar objek (perangkat daerah).
4. Menghitung kembali jarak dengan metode *Ward* menggunakan persamaan berikut:

$$I_{(AB)C} = \frac{(N_A + N_C)(I_{AC}) + N_B N_C(I_{BC}) - N_C(I_{AB})}{N_{AB} + N_C} \quad \dots (3)$$

   Keterangan:
   $I_{(AB)C}$: Jarak klaster *AB* dan *C*.
   I_{AC} : Jarak antara klaster *A* dan *C*.
   I_{BC} : Jarak antara klaster *B* dan *C*.
   N_A, N_B, N_C : Banyaknya objek pada klaster *A, B, C*.
5. Mengulangi tahap (2) hingga (4) hingga semua objek tergabung menjadi satu klaster.
6. Menentukan jumlah klaster optimalnya dengan menggunakan validasi klaster metode indeks *Dunn*, *Connectivity*, dan *Silhouette Coefficient* dengan masing-masong formulanya adalah sebagai berikut:
   - Indeks *Dunn*

$$DI(c) = \min_{1\le i\le n}\{min_{1\le j\le n}\left\{\frac{d(C_i, C_j)}{\max\limits_{1\le k,=n} d'(C_l)}\right\}\} \quad \dots (4)$$

Keterangan:
$i \neq j$: Pengamatan ke i tidak boleh sama dengan pengamatan ke j.
$d(C_i, C_j)$: Jarak antar klaster C_i dan C_j.
$d'(C_l)$: Jarak dalam klaster C_l.
n : Jumlah klaster.

- Indeks Connectivity

$$Conn(C) = \Sigma_{i=1}^{N} \Sigma_{j=1}^{L} X_{i,nn_{i(j)}} \quad \ldots (5)$$

Keterangan:
L : Parameter yang menentukan jumlah tetangga yang berkontribusi untuk nilai indeks
$nn_{i(j)}$: Pengamatan tetangga terdekat *i* ke *j*.

- Silhouette Coefficient

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i,j) \quad \ldots (6)$$

Keterangan:

$a(i)$: Perbedaan rata-rata objek *i* ke objek lainnya pada klaster A.

$d(i,j)$: Jarak objek ke *i* dengan objek ke *j*.

A : Klaster A.

$$d(i,C) = \frac{1}{|C|} \sum_{j \in C,} d(i,j) \quad \ldots (7)$$

$$b(i) = \min_{C \neq A} d(i,C) \quad \ldots (8)$$

Keterangan:
$d(i,C)$: Perbedaan rata-rata objek ke *i* dengan objek lain di klaster *C*.

C : Klaster C, dimana $C \neq A$.

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad \ldots (9)$$

$$SC = \max_{K} \frac{\sum_{i=1}^{n} S(i)}{n} \quad \ldots (10)$$

7. Melakukan pembentukan (penggabungan objek) klaster dengan metode *Ward.*

C. Hasil Penelitian dan Pembahasan

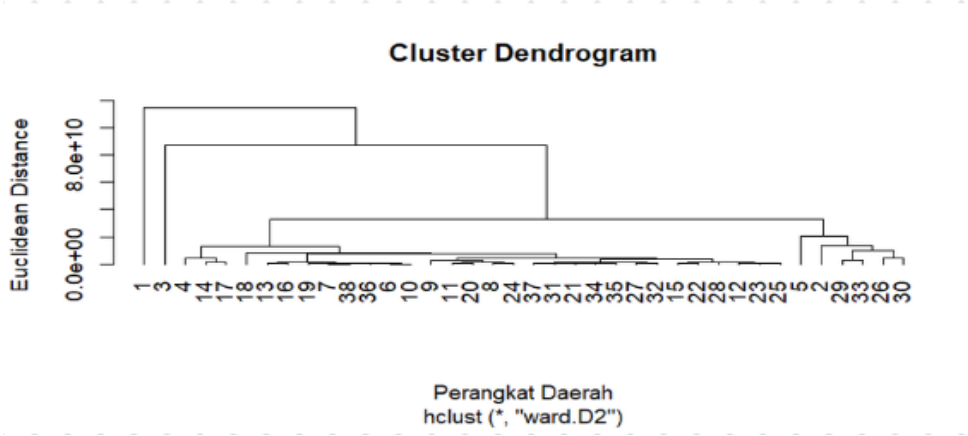
Berdasarkan data yang diolah diperoleh proses aglomerasi atau penggabungan dengan tanda negatif (-) pada angka di proses aglomerasi menandakan suatu objek sedangkan angka yang tidak memiliki tanda negatif menandakan suatu klaster yang terbentuk pada tahap tertentu.

Tabel 1. Proses Aglomerasi data

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 2	-6 -10	-7 -38	-13 -16	-34 -35	-36 1	-23 -25	-15 -22	-21 4	-31 8	2 5	-11 -20	-8 -24	-37 9
	14	15	16	17	18	19	20	21	22	23	24	25	26
1 2	-12 6	-19 10	-27 -32	-28 14	13 16	3 15	-14 -17	11 12	7 17	-9 21	-29 -33	18 22	-4 20
	27	28	29	30	31	32	33	34	35	36	37		
1 2	23 25	-26 -30	19 27	-18 29	24 28	26 30	-2 31	-5 33	32 34	-3 35	-1 36		

Pada Tabel 1 dapat diperhatikan bahwa proses aglomerasi dimulai dari penggabungan dua objek terdekat dan karena objek pada penelitian ini adalah Perangkat Daerah (PD_i), maka diperhatikan jarak antar perangkat daerahnya. Dimana pada tahap pertama objek ke 6 dan objek ke 10 yaitu PD_6 dan PD_{10} digabungkan. Dilanjutkan pada tahap ke 2, terjadi penggabungan antara objek ke 7 dan objek ke 38 yaitu PD_7 dan PD_8. Lalu, pada tahap ke 3 terdapat penggabungan antara objek ke 13 dan objek ke 16 yaitu PD_{13} dan PD_{16}. Kemudian, pada tahap ke 4 terjadi penggabungan antara objek ke 34 dan objek ke 35 yaitu PD_{34} dan PD_{35}. Langkah berikutnya, pada tahap ke 5 ini terlihat pada Tabel 1 bahwa terjadi penggabungan antara objek ke 36 dan klaster yang terbentuk pada tahap pertama yaitu PD_{36} dengan klaster yang beranggotakan PD_6 dan PD_{10}. Begitupun selanjutnya akan terus terjadi penggabungan antara dua objek, penggabungan antara objek dan suatu klaster, atau bahkan penggabungan antara dua klaster yang sudah terbentuk sebelumnya. Proses aglomerasi tersebut berhenti pada tahap 37 dimana terjadi penggabungan terakhir yaitu antara objek ke 1 dan klaster yang terbentuk pada tahap ke 36 yaitu PD_1 dengan klaster yang beranggotakan PD_3 beserta 36 perangkat daerah lainnya.

Karena proses aglomerasi telah berakhir, maka semua objek telah bergabung menjadi satu klaster. Pada proses ini dihasilkan beberapa klaster dengan masing-masing anggotanya yang dapat divisualisasikan sebagai berikut:



Gambar 1. Dendrogram Metode *Ward* Keseluruhan

Berdasarkan dendrogram pada Gambar 1, dendrogram tersebut menunjukkan jumlah kemungkinan klaster yang dapat terbentuk diantaranya. Kemudian pada dendrogram ini membentuk suatu cabang hierarki yang menggambarkan suatu jarak atau perbedaan dimana semakin dekat jarak dua buah objek, maka objek tersebut semakin mirip karakteristiknya satu sama lain. Pada dendrogram tersebut terlihat bahwa PD_6 dan PD_{10} merupakan objek pertama yang tergabung menjadi satu klaster karena pada dendrogram letak kedua objek berada di sub cabang yang paling bawah sehingga menjelaskan bahwa kedua objek tersebut mempunyai jarak yang terdekat atau mempunyai karakteristik yang termirip satu sama lain. Kemudian dimisalkan kembali pada dendrogram terlihat bahwa PD_{26} dan PD_{30} mempunyai jarak yang saling berdekatan dimana hal ini menjelaskan bahwa kedua perangkat daerah memiliki karakteristik yang mirip, tetapi PD_{30} akan memiliki karakteristik yang jauh berbeda dengan PD_1 karena perangkat daerah tersebut terletak jauh di cabang yang berbeda. Maka, berdasarkan hal tersebut menunjukkan bahwa posisi objek yang saling bersebelahan pada dendrogram memberikan suatu informasi bahwa jarak yang berdekatan memungkinkan suatu objek tergabung dalam satu klaster yang sama. Maka dari itu dapat dikatakan bahwa semakin dekat jarak antar objek semakin mirip karakteristiknya, sebaliknya semakin jauh jarak antar objek semakin berbeda karakteristiknya. Dan dalam pemotongan cabang ke arah atas, maka semakin heterogen anggota klasternya.

Analisis klaster menunjukkan anggota dari setiap klaster untuk banyaknya klaster tertentu bukan memutuskan jumlah klaster yang terbentuk secara subjektif. Maka dari itu,

diperlukan validasi klaster untuk menentukan jumlah klaster terbaik. Penentuan klaster terbaik ini akan menggunakan validasi klaster dengan validasi Dunn index dan Silhoutte coefficient yang memiliki ketentuan bahwa klaster yang mempunyai nilai Dunn index atau nilai Silhouette Coefficient tertinggi serta nilai Connectivity index terendah merupakan klaster yang baik digunakan. Berikut hasil dari nilai validasi masing-masing klaster yang terbentuk.

Tabel 2. Penentuan Klaster Terbaik Metode *Ward*

Klaster	Connectivity Index (CI)	Dunn Index (DI)	Silhouette Coefficient (SC)
2	4,9659	0,3422	0,7559
3	11,9647	0,0568	0,4293
4	12,5758	0,0654	0,4072
5	17,6463	0,0712	0,4471

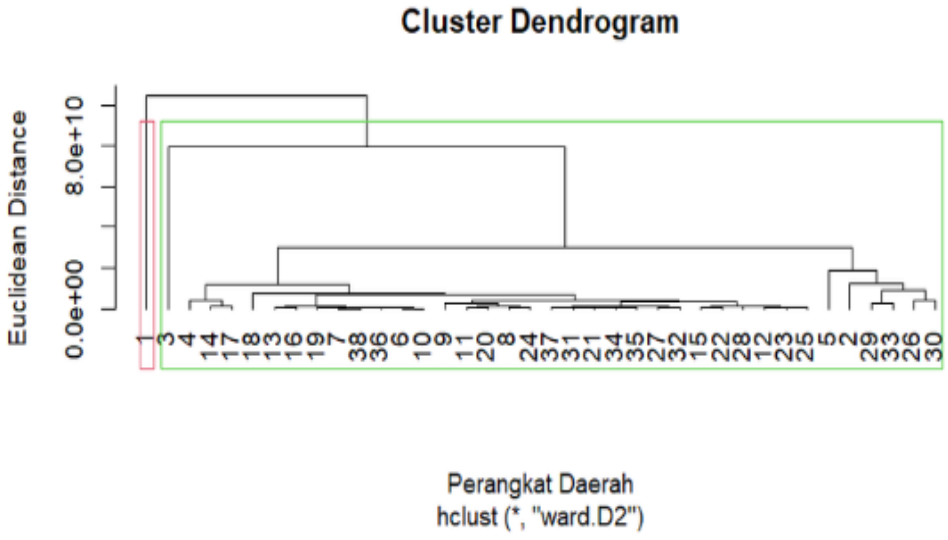
Berdasarkan Tabel 2 diperoleh bahwa pada pembentukan 2 klaster diperoleh nilai indeks *Dunn* dan indeks *Silhouette* tertinggi serta nilai indeks *Connectivity* terendah, sehingga pada metode *Ward* akan dilakukan pengelompokkan perangkat daerah dengan jumlah klaster yang terbentuk adalah 2 klaster. Hasil pengelompokkan perangkat daerah dengan 2 klaster disajikan sebagai berikut:

Tabel 3. Anggota Klaster untuk Metode Ward

Anggota Klaster		
	Klaster 1	Klaster 2
Perangkat Daerah	Perangkat Daerah 1 (PD_1)	Perangkat Daerah Lainnya

Berdasarkan 2 diatas dapat disimpulkan bahwa setiap klaster mempunyai anggotanya masing-masing yaitu pada klaster 1 beranggotakan satu perangkat daerah yaitu PD_1 dan pada klaster 2 beranggotakan tiga puluh tujuh perangkat daerah yaitu PD_2, PD_3, PD_4, dan 34 perangkat daerah lainnya.

Adapun visualisasi yang terbentuk untuk metode *ward* dengan jumlah klaster terbaik sebanyak 2 klaster berdasarkan Tabel 2, dapat ditunjukkan pada dendrogram sebagai berikut:



Gambar 2. Dendrogram Metode *Ward* dengan K = 2

Pada Gambar 2 diatas menunjukkan bahwa anggota klaster 1 pada kotak berwarna merah yaitu PD_1 dan anggota klaster 2 pada kotak berwarna hijau yaitu PD_2, PD_3, PD_4, dan 34 perangkat daerah lainnya.

Berdasarkan dendrogram kotak berwarna merah pada dendrogram menunjukkan PD_1 sebagai objek yang mempunyai kriteria yang berbeda sehingga perangkat daerah tersebut dipisahkan dalam klaster yang berbeda tidak tergabung dengan perangkat daerah lainnya di kotak berwarna hijau.

Selanjutnya dilakukan perhitungan rata-rata variabel dalam setiap klasternya untuk memahami karakteristik masing-masing klaster dan diketahui klaster 1 mempunyai rata-rata yang lebih besar di seluruh variabel (jenis pajak) dibandingkan dengan klaster 2. Oleh karena itu, dapat dikatakan bahwa klaster 1 beranggotakan perangkat daerah yang mempunyai tarikan pajak yang tinggi di SIPD sedangkan klaster 2 beranggotakan perangkat daerah dengan tarikan pajak rendah.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Hasil dari proses pengelompokan perangkat daerah berdasarkan data total rekapitulasi tarikan pajak di Sistem Informasi Pemerintah Daerah (SIPD) pada bulan Januari-Agustus tahun 2023 dengan menggunakan analisis klaster hierarki metode ward diperoleh sebanyak 2 klaster dengan jumlah anggota klaster 1 sebanyak 1 Perangkat Daerah dan jumlah anggota klaster 2 sebanyak 37 perangkat daerah.
2. Dapat disimpulkan, penelitian ini memberikan hasil akhir pengelompokan yakni Perangkat Daerah 1 (PD_1) sebagai satu-satunya anggota dalam klaster 1 dan Perangkat Daerah 2 (PD_2) serta 36 perangkat daerah lainnya berada dalam klaster 2. Nilai rata-rata tiap jenis pajak yang diberikan oleh klaster 1 lebih besar dibandingkan nilai rata-rata tiap jenis pajak yang diberikan oleh klaster 2, sehingga dapat dikatakan bahwa klaster 1 memiliki tarikan pajak paling besar.

Acknowledge

Puji dan Syukur penulis panjatkan kehadirat Allah Subhanahu Wa Ta'ala atas limpahan rahmat dan karunia-Nya penulis dapat menyelesaikan jurnal artikel ini. Shalawat serta salam semoga selalu tercurah limpahkan kepada Nabi Muhammad SAW juga kepada keluarganya, para sahabatnya dan pengikutnya hingga akhir zaman. Penulis menyadari banyak pihak yang memberikan dukungan dan bantun dalam menyelesaikan artikel ini diantaranya Bapak Abdul Kudus, S.Si, M.Si., Ph.D., selaku ketua prodi Statistika dan Bapak Suwanda, selaku dosen pembimbing pada pembuatan jurnal artikel ini. Dengan adanya laporan ini diharapkan dapat bermanfaat bagi para pembaca dan dapat dijadikan referensi demi pengembangan ke arah yang lebih baik. Dalam penulisan laporan ini penulis menyadari masih terdapat kekurangan, penulis memohon maaf dan mengharapkan adanya kritik dan saran sebagai bahan evaluasi untuk penyempurnaan laporan ini.

Daftar Pustaka

[1] Laeli, Sofya. (2014), Analisis Cluster dengan *Average Linkage Method* dan *Ward's Method* untuk Data Responden Nasabah Asuransi Jiwa Unit Link. Skripsi, Universitas Negeri Yogyakarta.

[2] Zulfa, Farida, B2A218004 (2019), Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan Menggunakan *K-Means* dan *K-Medoids.* Sarjana / Sarjana Terapan (S1/D4) tesis.

[3] Veriani, Rafika. (2020). Analisis Cluster Dalam Pengelompokkan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode *Complete Linkage*, *Average Linkage* dan *Ward*. Skripsi tidak dipublikasikan. Program Sarjana, Program Studi Matematika, Universitas Islam Negeri Sunan Ampel Surabaya.

[4] Veriani, Rafika. (2020). Analisis *Cluster* Dalam Pengelompokkan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode *Complete Linkage*,

Average Linkage, dan *Ward*. Program Sarjana, Program Studi Matematika, Universitas Islam Negeri Sunan Ampel Surabaya.

[5] August 12, 2021). Pajak Penghasilan Pasal 21. KPPN Kotabumi. Retrieved September 23, 2023, from https://djpb.kemenkeu.go.id/kppn/kotabumi/id/informasi/perpajakan/pph-pasal-21.html

[6] PPh Pasal 22: Tarif, Cara Hitung dan Lapor SPT Masa PPh 22. (May 14, 2023). Mekari klikpajak. Retrieved September 23, 2023, from https://klikpajak.id/blog/pph-pasal-22-dan-lapor-spt-pph-22/.)

[7] Pengertian PPh Pasal 23. HiPajak. Retrieved September 23, 2023, from https://www.hipajak.id/artikel-pengertian-pph-pasal-23.

[8] PPh Pasal 4 Ayat 2 (Pajak Penghasilan Pasal 4 Ayat 2). (February 21, 2023). OnlinePajak. Retrieved September 23, 2023, from https://www.onlinepajak.com/seputar-pajakpay/pajak-penghasilan-pph-pasal-4-ayat-2-a.)

[9] Arisa, J. W., & Hajarisman, N. (2023). Determinasi Derajat Kelangsungan Hidup Anak Menggunakan Multigroup Structural Equation Modeling. *Jurnal Riset Statistika*, *3*(2), 147–154. https://doi.org/10.29313/jrs.v3i2.3047

[10] Bernardino, M. (2023). Penerapan CUSUM-Tukey's Control Chart untuk Mendeteksi Perubahan Rata-Rata Proses pada Data Non-Normal. *Jurnal Riset Statistika*, *3*(2), 119–124. https://doi.org/10.29313/jrs.v3i2.2955

[11] Fatmawati, & Rifai, N. A. K. (2023). Klasifikasi Penyakit Diabetes Retinopati Menggunakan Support Vector Machine dengan Algoritma Grid Search Cross-validation. *Jurnal Riset Statistika*, 79–86. https://doi.org/10.29313/jrs.v3i1.1945