

Penerapan Metode *K-Means* Untuk Memprediksi Pola Kedatangan Penduduk di Kota Jakarta

Dhimas Hernandi^{1*}, Arief Jananto²

Program Studi Sistem Informasi, Universitas Stikubank Semarang, Semarang, Indonesia

**e-mail Corresponding Author: dhimashernandi@gmail.com*

Abstract

Population density in an area has the potential to provoke complex problems, the arrival of people from one area to another in large numbers also has the potential to cause unequal distribution of population on an island. The city of Jakarta in January to May 2021 recorded the arrival of a population of 52,309, the impact of this large number of arrivals will greatly affect the quality of life and the density of mobility in an area. This article examines the use of the K-Means Clustering algorithm to classify the number of arrivals to the city of Jakarta by urban villages. 267 urban villages in the city of Jakarta on population arrival data during January to May 2021 implemented in the K-Means algorithm to determine 3 groups of average number of arrivals (low, medium and high). Each group will represent the pattern of arrivals in each month based on the average arrival and the average in 5 months from January to May 2021. The test results found a increase in arrivals in March in all groups.

Kata kunci: *Population; Arrival; Urban Villages; K-Means Algorithm*

Abstrak

Kepadatan penduduk dalam suatu daerah berpotensi menimbulkan masalah yang kompleks. Kedatangan penduduk dari satu daerah ke daerah yang lain dalam jumlah yang besar juga berpotensi terhadap timpangnya sebaran penduduk di suatu pulau. Kota Jakarta pada bulan Januari hingga Mei tahun 2021 tercatat kedatangan penduduk sebanyak 52.309, dampak dari kedatangan dalam jumlah besar ini akan sangat berpengaruh terhadap kualitas hidup dan padatnya mobilitas pada suatu daerah. Artikel ini menguji penggunaan algoritme *K-Means Clustering* untuk mengelompokkan jumlah kedatangan penduduk ke kota Jakarta berdasarkan kelurahan. 267 kelurahan di kota Jakarta pada data kedatangan penduduk selama Januari hingga Mei 2021 diimplementasikan pada algoritme *K-Means* untuk menentukan 3 kelompok rata-rata jumlah pendatang (rendah, menengah dan tinggi). Setiap kelompok akan merepresentasikan pola kedatangan di tiap bulannya berdasarkan rata-rata kedatangan dan rata-rata dalam 5 bulan dari Januari hingga Mei 2021. Hasil pengujian mendapati lonjakan kedatangan pada bulan Maret pada semua kelompok.

Kata kunci: *Sebaran penduduk; Kedatangan; Kelurahan; Algoritme K-Means Clustering*

1. Pendahuluan

Pada tahun 2020 situs <https://data.jakarta.go.id> mempublikasikan bahwa penduduk di Jakarta berjumlah 11.100.929 jiwa menghuni daratan dengan luas sekitar 664,01 km². Kedatangan penduduk ke suatu daerah memancing pertumbuhan ekonomi karena mobilitas akan semakin tinggi tetapi efek dari kepadatan penduduk juga harus diperhatikan. Dari 267 kelurahan di Jakarta kedatangan penduduk mempunyai jumlah yang beragam, terhitung dari Januari hingga Mei 2021 total kedatangan penduduk ke Jakarta mencapai 52.309 penduduk. Berangkat dari latar belakang kedatangan tersebut Pemerintah Kota Jakarta harus mempertimbangkan masalah yang berpotensi muncul seiring dengan kedatangan penduduk.

Faktor yang berpengaruh terhadap pengambilan kebijakan pada daerah dengan intensitas kedatangan penduduk tinggi erat kaitannya dengan prediksi kedatangan di waktu mendatang. Sebaran penduduk yang terpusat pada suatu daerah berdampak pada tingkat keefektifan fasilitas tatanan tempat tinggal, sanitasi, dan juga mobilitas yang tinggi. Langkah

untuk menganalisa banyaknya penduduk yang datang salah satunya dengan menggunakan Metode *clustering* dengan mengelompokkan daerah dengan pengunjung dengan skala rendah, menengah dan tinggi agar pemerintah mempunyai acuan untuk melakukan kebijakan untukantisipasi kepadatan penduduk.

Metode *clustering* merupakan metode analisa data yang dapat digunakan dalam memecahkan masalah dalam pengelompokan data [1]. Metode *clustering* termasuk dalam kajian ilmu data mining dimana data mining adalah proses ekstraksi informasi data berukuran besar, sedangkan *clustering* berguna untuk menemukan pola distribusi di dalam sebuah dataset yang berguna untuk proses analisa data [2]. Salah satu algoritme yang bisa dipakai dalam metode *clustering* adalah *K-Means* karena algoritme tersebut merupakan algoritme yang bisa dipakai untuk data dengan ukuran yang besar dengan waktu relatif cepat dan efisien. *Clustering* banyak digunakan dalam penggalan pola kecenderungan data yang besar, perkembangan ukuran data yang pesat seperti data sebaran pandemi *Covid-19* [3], [4], data yang berhubungan dengan perpindahan penduduk [5]–[8], penanggulangan bencana [2], fasilitas kesehatan [1], fasilitas masyarakat [9] dan bidang akademik [10]–[12]. Dalam artikel ini data sebaran akan diimplementasikan dalam algoritme *K-Means* dengan harapan menemukan pola kedatangan penduduk ke Kota Jakarta.

2. Tinjauan Pustaka

Data mining merupakan proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer secara otomatis untuk menganalisa dan mengekstraksi pengetahuan [13]. Penelitian M Fauzan, Windarto Perdana W dan Irmanita Nasution menerapkan algoritma untuk mengelompokkan penduduk miskin, data penduduk yang besar akan di ekstraksi menjadi data yang lebih kecil dengan teknik *clustering*. *Clustering* adalah algoritma pengelompokan data dengan jumlah besar diproses menjadi bagian kecil dengan atribut kemiripan tertentu [14]. Tujuan dari analisa didasarkan untuk membagi wilayah menjadi beberapa kelompok yang bertujuan untuk memberi masukan kepada pemerintah agar dapat menangani penyebaran bantuan untuk daerah tingkat penduduk miskin yang berdampak pada pertumbuhan ekonomi di Indonesia.

Penelitian pada pemetaan kepadatan penduduk berdasarkan jumlah penduduk di Kota Medan [15] membagi kelompok berdasarkan kelurahan, dari penelitian yang dilakukan menghasilkan 3 *cluster*: *cluster 1* untuk kelurahan yang sangat padat penduduk, *cluster 2* untuk daerah dengan penduduk padat, dan *cluster 3* untuk kelurahan dengan kepadatan penduduk sedang. *Cluster 1* berisi 121 kelurahan yang tersebar di 21 kecamatan, *cluster 2* berisi 30 kecamatan dan tersebar di 10 kecamatan, sementara untuk *cluster 3* tidak lagi ditemukan daerah dengan tingkat penduduk sedang.

Berlatar belakang pada masalah kepadatan penduduk pada kabupaten Simalungun peneliti menerapkan algoritme *K-Means Clustering* untuk membagi kelompok berdasarkan kecamatan yang dikelompokkan menjadi 3 cluster dengan jumlah kepadatan penduduk rendah, sedang dan tinggi [5]. Terdapat 4 kecamatan yang masuk kedalam *cluster 1*, 11 kecamatan yang masuk dalam *cluster 2* dan 17 kecamatan yang masuk kedalam *cluster 3*. Total *record* yang dipakai sebanyak 32.

Potensi transmigrasi menjadi penelitian [16] pada studi kasus Kabupaten Malang. Parameter yang dipakai adalah jumlah penduduk, laju pertumbuhan penduduk, jumlah usia subur, dan kepadatan penduduk. jika jumlah penduduk dan laju pertumbuhan penduduk berbanding lurus maka ada potensi penduduk melakukan transmigrasi. Dari hasil clustering mendapatkan *cluster 1* terdapat 4 kecamatan yang tidak berpotensi terjadi transmigrasi, 27 kecamatan yang masuk dalam *cluster 2* yang berpotensi terjadi transmigrasi dan 2 kecamatan yang sangat berpotensi terjadi transmigrasi.

Penelitian mengenai pengelompokan kesejahteraan rakyat di kabupaten Karawang [17], menggali informasi mengenai daerah mana saja yang memiliki tingkat kesejahteraan tingkat rendah sehingga pemerintah dapat memprioritaskan kesejahteraan rakyat pada wilayah tersebut. Pengelompokan menghasilkan 3 *cluster*, *cluster 1* merupakan *cluster* dengan tingkat kesejahteraan rakyat tinggi terdiri dari 7 kecamatan, *cluster 2* merupakan *cluster* dengan tingkat kesejahteraan rakyat sedang terdiri dari 8 kecamatan dan *cluster 3* merupakan *cluster* dengan tingkat kesejahteraan rendah terdiri dari 7 kecamatan.

Dari beberapa risetterdahulu yang ditinjau, penggunaan algoritme *K-Means* merupakan algoritma yang secara maksimal bisa digunakan untuk waktu yang relatif cepat, pengkategorian

berdasarkan kesamaan objek berdasarkan jarak yang ada pada setiap atribut membuat klasterisasi memiliki karakteristik yang jelas berbeda antar klaster, metode penelitian menggunakan KDD (*Knowledge Discovery in Database*) proses pengolahan data dilakukan dengan RStudio. Diharapkan dengan metode ini mendapatkan pola kedatangan yang berpengaruh terhadap kemungkinan kepadatan penduduk di Jakarta berdasarkan kelurahan dengan akurasi yang baik.

3. Metodologi

3.1 Algoritme K-Means

Pengelompokan dilakukan dengan perhitungan *Euclidean Distance* dimana pencarian anggota *cluster* berdasarkan jarak terdekat dari batas-batas *cluster*. Batas *cluster* k ditetapkan pada awal dan setiap data dihitung secara iteratif hingga data terpartisi sesuai dengan rata-rata tiap titik pusat *cluster* k. Algoritme *K-Means* dijelaskan sesuai tahapan berikut:

1. Menentukan jumlah cluster (k).
2. Menentukan titik pusat *cluster* (k) secara acak.
3. Hitung jarak terhadap titik *centroid* dengan persamaan

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots(1)$$

Keterangan:

d = jarak

x = koordinat objek

y = koordinat titik pusat *cluster*

n = banyaknya objek

4. Kelompokkan hasil penghitungan ke dalam *cluster* berdasarkan jarak terdekat terhadap titik pusat *cluster*.
5. Tentukan titik pusat *cluster* yang baru dengan menghitung rata-rata anggota tiap *cluster*.

$$c = \frac{\sum m}{n} \dots\dots\dots(2)$$

Keterangan:

C = titik pusat *cluster*

$\sum m$ = total anggota *cluster*

n = banyak *cluster*

6. Lakukan perhitungan ulang seperti langkah 3 terhadap titik pusat *cluster* yang baru hingga nilai rata-rata titik pusat *cluster* tidak berubah.

3.2 Objek Penelitian

Dataset yang diimplementasikan pada algoritme *K-Means Clustering* pada penelitian ini menggunakan dataset kedatangan penduduk dari luar Jakarta ke Jakarta tahun 2021 yang diunduh dari lama <https://data.jakarta.go.id> dari bulan Januari sampai dengan Mei. Dataset dikelompokkan berdasarkan kedatangan rendah, menengah dan tinggi. Data yang dipakai sejumlah 267 kelurahan dan hasil dari pengelompokkan akan merepresentasikan rata-rata kedatangan, titik puncak kedatangan pada masing-masing kelompok dan perbandingan luas wilayah sebagai tujuan kedatangan.

4. Hasil dan Pembahasan

4.1 Data Pre-Processing

Peneliti menggunakan dataset kedatangan penduduk ke Jakarta dari bulan Januari sampai dengan Mei 2021, data terdiri dari 2670 *record* dan selanjutnya di *pre-processing* dengan data selection untuk pemilihan data yang akan diproses.

Tabel 1. Potongan Tabel Data Sebelum di Seleksi

Bulan	Kota Kabupaten	Kecamatan	Kelurahan	Jenis Kelamin	Jumlah
1	Adm. Kepulauan Seribu	Kepulauan Seribu Utara	Pulau Panggang	Laki-Laki	7
1	Adm. Kepulauan Seribu	Kepulauan Seribu Utara	Pulau Kelapa	Laki-Laki	0
1	Adm. Kepulauan Seribu	Kepulauan Seribu Utara	Pulau Harapan	Laki-Laki	0
1	Adm. Kepulauan Seribu	Kepulauan Seribu Selatan	Pulau Untung Jawa	Laki-Laki	1
1	Jakarta Pusat	Tanah Abang	Kebon Kacang	Laki-Laki	11
1	Jakarta Pusat	Tanah Abang	Kampung Bali	Laki-Laki	3
1	Jakarta Pusat	Johar Baru	Johar Baru	Laki-Laki	17
1	Jakarta Pusat	Johar Baru	Kampung Rawa	Laki-Laki	9

Kemudian data diseleksi dengan melakukan transpose tabel, merubah bulan menjadi atribut dan melakukan penjumlahan data laki-laki serta perempuan untuk menampilkan data jumlah sebaran secara keseluruhan. Data setelah diseleksi pada tabel 2.

Tabel 2. Potongan Tabel Data Setelah Diseleksi

No	Kelurahan	Januari	Februari	Maret	April	Mei
1	Pulau Panggang	12	3	7	5	0
2	Pulau Kelapa	0	2	3	11	1
3	Pulau Harapan	0	0	2	8	1
4	Pulau Untung Jawa	1	5	4	3	4
5	Pulau Tidung	7	1	2	1	2
6	Pulau Pari	6	0	4	7	0
7	Gambir	8	2	7	10	2
8	Cideng	6	13	19	21	8
...
...
256	Cibubur	83	98	79	88	50
257	Kelapa Dua Wetan	52	79	100	70	36
258	Susukan	52	58	22	55	35
259	Rambutan	51	33	54	74	23
260	Cipayung	43	20	47	30	36
261	Cilangkap	49	45	44	40	39
262	Pondok Ranggon	43	33	44	54	22
263	Munjul	28	19	54	35	20
264	Setu	33	38	74	38	20
265	Bambu Apus	31	34	56	30	43
266	Lubang Buaya	71	75	93	99	64
267	Ceger	22	14	36	21	10

4.2 Perhitungan Manual Algoritme *K-Means*

1. Menentukan jumlah *cluster* (k)
Penelitian ini menggunakan 3 *cluster* untuk mengelompokkan kelurahan berdasarkan *cluster* rendah, menengah dan tinggi.
2. Menentukan titik pusat *cluster* (k) secara acak.

Pusat *cluster* iterasi 1 disajikan dalam Tabel 3.

Tabel 3. Titik Pusat Cluster Iterasi 1

Cluster	Januari	Februari	Maret	April	Mei
Cluster 1	0	2	3	11	1
Cluster 2	16	10	10	21	4
Cluster 3	36	15	32	26	7

3. Hitung jarak terhadap titik pusat *cluster*.

Perhitungan jarak terhadap titik pusat iterasi 1 adalah sebagai berikut:

$$d(1,1) = \sqrt{(0 - 12)^2 + (2 - 2)^2 + (3 - 7)^2 + (11 - 5)^2 + (1 - 0)^2}$$

$$d(1,1) = 14.07124728$$

$$d(1,2) = \sqrt{(0 - 0)^2 + (2 - 2)^2 + (3 - 3)^2 + (11 - 11)^2 + (1 - 1)^2}$$

$$d(1,2) = 0$$

$$d(1,3) = \sqrt{(0 - 0)^2 + (2 - 0)^2 + (3 - 2)^2 + (11 - 8)^2 + (1 - 1)^2}$$

$$d(1,3) = 3.741657387$$

Perhitungan dilanjutkan hingga data ke 267 dan juga untuk titik pusat *cluster 2* dan *cluster 3*.

4. Kelompokkan hasil penghitungan ke dalam *cluster* berdasarkan jarak terdekat terhadap titik pusat *cluster*.

Mengelompokkan data ke dalam masing-masing *cluster* berdasarkan jarak terdekat terhadap pusat titik *cluster*. Hasil pengelompokan data pada hasil iterasi pertama tertera pada tabel 4.

Tabel 4. Potongan Tabel Hasil Perhitungan Pertama

No	Kelurahan	Januari	Februari	Maret	April	Mei	Hasil C1	Hasil C2	Hasil C3	Cluster
1	Pulau Panggang	12	3	7	5	0	14.07124728	18.60107524	42.83689998	1
2	Pulau Kelapa	0	2	3	11	1	0	21.86321111	50.66557016	1
3	Pulau Harapan	0	0	2	8	1	3.741657387	24.45403852	52.73518749	1
4	Pulau Untung Jawa	1	5	4	3	4	9.16515139	24.69817807	51.44900388	1
5	Pulau Tidung	7	1	2	1	2	12.32882801	25.0998008	50.8625599	1
6	Pulau Pari	6	0	4	7	0	7.615773106	21.16601049	48.15599651	1
7	Gambir	8	2	7	10	2	9.055385138	16.18641406	43.11612227	1
8	Cideng	6	13	19	21	8	23.70653918	14.35270009	33.15116891	2
...
...
257	Kelapa Dua Wetan	52	79	100	70	36	150.8244012	132.5971342	108.411254	3
258	Susukan	52	58	22	55	35	96.40020747	76.55716818	61.88699379	3
259	Rambutan	51	33	54	74	23	103.033975	82.82511696	59.94163828	3
260	Cipayung	43	20	47	30	36	75.46522378	57.47173218	34	3
261	Cilangkap	49	45	44	40	39	90.64215355	71.10555534	49.32544982	3
262	Pondok Ranggon	43	33	44	54	22	82.3468275	61.86275131	39.06404997	3
263	Munjul	28	19	54	35	20	67.90434449	51.11751168	28.53068524	3

No	Kelurahan	Januari	Februari	Maret	April	Mei	Hasil C1	Hasil C2	Hasil C3	Cluster
264	Setu	33	38	74	38	20	92.28217596	75.59100476	51.1370707	3
265	Bambu Apus	31	34	56	30	43	83.18052657	67.22350779	47.68647607	3
266	Lubang Buaya	71	75	93	99	64	173.7325531	154.3470116	130.8587024	3
267	Ceger	22	14	36	21	10	43.56604182	27.64054992	15.71623365	3

5. Tentukan titik pusat cluster yang baru dengan menghitung rata-rata anggota tiap *cluster*. Setelah data dikelompokkan ke dalam masing-masing *cluster*, titik pusat *cluster* dihitung berdasarkan rata-rata anggota *cluster* yang selanjutnya akan didapatkan titik pusat *cluster* yang baru.

Cluster	Januari	Februari	Maret	April	Mei
Cluster 1	5.366667	4.4	6.4	6.2	3.333333
Cluster 2	19.65556	21.2	27.01111	23.77778	14.2
Cluster 3	54.14966	57.87755	70.91156	63.32653	39.53061

Lanjutkan perhitungan hingga rata-rata titik pusat *cluster* pada iterasi tidak ada perubahan, jika sudah tidak terdapat perubahan titik pusat *cluster* maka hasil dari pengelompokan kelurahan berdasarkan *cluster* sudah selesai.

Titik pusat *cluster* akhir yang terbentuk pada iterasi ke-20 adalah sebagai berikut:

Cluster	Januari	Februari	Maret	April	Mei
Cluster 1	17.78676	18.73529	24.11765	21.52206	12.96324
Cluster 2	49.59259	51.76852	64.18519	56.71296	35.26852
Cluster 3	91.95652	104.7391	123.2609	112.3043	70.30435

Pada tabel 5 disajikan data hasil perhitungan data setelah stabil dan rata-rata titik pusat *cluster* tidak berubah:

Tabel 5. Hasil Perhitungan Pengelompokan Kelurahan

Cluster	Titik Pusat Cluster Januari	Titik Pusat Cluster Februari	Titik Pusat Cluster Maret	Titik Pusat Cluster April	Titik Pusat Cluster Mei	Banyaknya Anggota
Cluster 1	17.78676	18.73529	24.11765	21.52206	12.96324	136
Cluster 2	49.59259	51.76852	64.18519	56.71296	35.26852	108
Cluster 3	91.95652	104.7391	123.2609	112.3043	70.30435	23

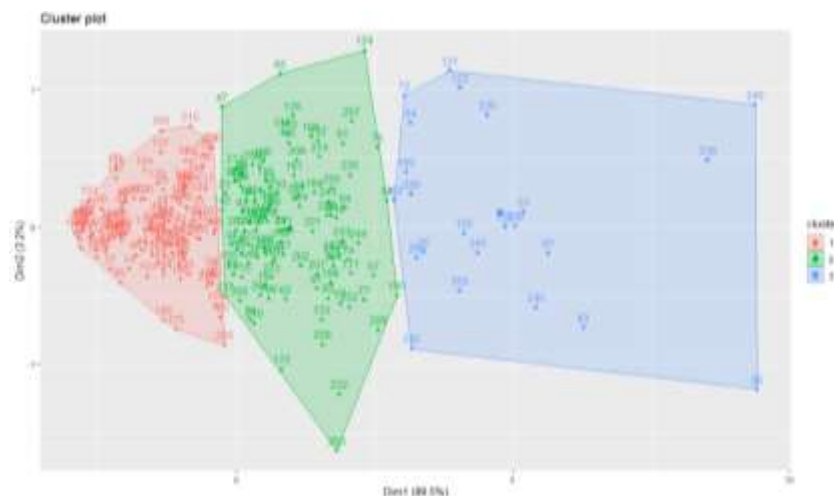
4.3 Implementasi Menggunakan RStudio

1. Visualisasi Data dengan Datatable

Untuk melihat data didalam datatable jalankan *script* berikut:

```
Datatable (kedatangan_data, caption = "Data Kedatangan
Penduduk ke Wilayah Jakarta")
```

Hasil visualisasi ditampilkan dalam gambar 1.



Gambar 3. Hasil Visualisasi Plot

Dari visualisasi plot gambar 4.8 didapatkan penjelasan sebagai berikut:

1. *Cluster 1* ditunjukkan dengan warna merah, anggota *cluster* sebanyak 136 kelurahan. Rata-rata untuk tiap clusternya adalah:

Januari	: 17.78676
Feburari	: 18.73529
Maret	: 24.11765
April	: 21.52206
Mei	: 12.96324
2. *Cluster 2* ditunjukkan dengan warna hijau, anggota *cluster* sebanyak 108 kelurahan dengan titik pusat *cluster*:

Januari	: 49.59259
Feburari	: 51.76852
Maret	: 64.18519
April	: 56.71296
Mei	: 35.26852
3. *Cluster 3* ditunjukkan dengan warna biru dengan anggota *cluster* sebanyak 23 kelurahan, dengan titik pusat *cluster*:

Januari	: 91.95652
Feburari	: 104.73913
Maret	: 123.26087
April	: 112.30435
Mei	: 70.30435

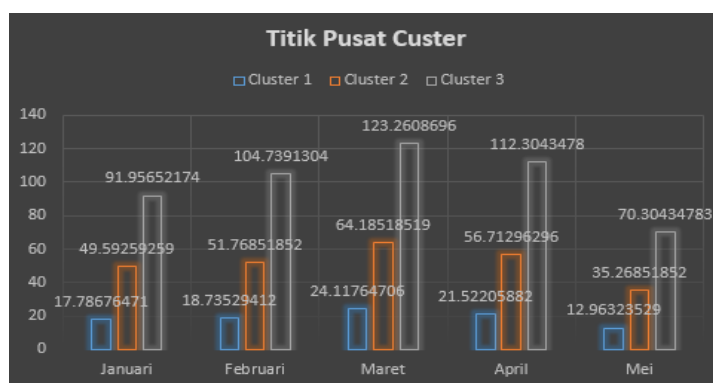
4.4 Interpretasi Hasil *Cluster*

Total pendatang selama Januari hingga Mei 2021 adalah 52.309 penduduk. Terdapat 136 kelurahan yang masuk dalam kategori *cluster 1*, dalam cluster ini kelurahan memiliki tingkat pendatang antara 12 hingga 24, rata-rata sebanyak 20 penduduk di tiap bulan selama 5 bulan antara Januari hingga Mei 2021. Jumlah total kedatangan untuk *cluster 1* sebanyak 12.937 atau 25% dari angka total kedatangan.

Sementara untuk kelurahan yang masuk kedalam *cluster 2* terdapat 108 kelurahan *cluster* ini memiliki tingkat pendatang antara 36 hingga 65 penduduk atau di rata-rata sebanyak 52 penduduk di tiap bulannya. Jumlah total kedatangan pada *cluster 2* ini sebanyak 27.813 penduduk prosentase 53% dari total kedatangan.

Pada *cluster* ketiga dengan tingkat kedatangan tinggi terdapat 23 kelurahan, range angka kedatangan penduduk ini berkisar antara 70 hingga 124 penduduk atau rata-rata 109 penduduk di antara bulan Januari hingga Mei 2021. Jumlah total kedatangan untuk *cluster 3* ini sebanyak 11.559 penduduk dengan prosentase 22% dari total kedatangan.

Sesuai dengan tujuan awal penelitian untuk menemukan pola kedatangan di Kota Jakarta maka dari hasil yang didapatkan bahwa *cluster 3* memiliki pendatang yang paling banyak, indikasi ini mengarah kepada ketimpangan pemerataan dan perlu dibandingkan dengan ketersediaan luas area yang dihuni. Terdapat potensi kepadatan karena banyak kelurahan anggota *cluster 3* memiliki luas wilayah yang relatif kecil dibanding dengan luas wilayah pada anggota *cluster 1* dan *cluster 2*. *cluster 3* memiliki anggota kelurahan tersebut diantaranya Penjaringan (3.95 m²), Sukapura (5.61 m²), Kalibaru (2.47 m²), Duri Kosambi (5.91 m²), Rawa Buaya (4.07 m²), Kapuk (5.63 m²), Cengkareng Timur (4.52 m²), Kalideres (5.71 m²), Semanan (5.98 m²), Tegal Alur (4.97 m²), Pegadungan (8.67 m²), Pejaten Timur (2.68 m²), Jagakarsa (4.85 m²), Jatinegara (6.6 m²), Penggilingan (4.48 m²), Cakung Timur (9.81 m²), Pulo Gebang (6.92 m²), Cakung Barat (6.12 m²), Klender (3.05 m²), Pondok Kelapa (5.72 m²), Ciracas (3.93 m²), Cibubur (4.5 m²) dan Lubang Buaya (3.72 m²).



Gambar 4. Hasil Perhitungan K-Means Clustering

Dalam menghadapi setiap lonjakan pada daerah kedatangan Pemerintah Republik Indonesia harus lebih giat dalam melakukan upaya transmigrasi [16] ke daerah yang lain untuk pemerataan penduduk yang bisa memancing ketertarikan penduduk di daerah yang masih belum berkembang untuk pemerataan di Indonesia. Banyak masalah yang mengakibatkan perpindahan dalam suatu wilayah ke daerah yang lebih potensial dalam perekonomian karena perspektif lebih mudah mengubah status sosial bagi rakyat miskin [17]. Dalam penelitian pertumbuhan penduduk perkotaan dan perkembangan pola distribusinya pada kawasan metropolitan secara jelas memberikan gambaran urbanisasi beserta aktivitasnya menyebabkan kepadatan lebih tinggi daripada kawasan-kawasan lain disekitarnya [18].

5. Simpulan

Kedatangan penduduk dalam 5 bulan pada tahun 2021 memberikan gambaran puncak kedatangan pada bulan maret pada semua kelurahan. Penggunaan klasterisasi pada penelitian memberikan informasi kepada pemerintah terkait pemerataan penduduk, pemenuhan kelancaran mobilitas untuk transportasi, persiapan menghadapi isu-isu yang timbul yang berefek pada pembenahan kelayakan penghidupan karena banyaknya penduduk yang datang ke suatu wilayah tetapi tidak berbanding lurus dengan luas area tempat tinggal.

Saran untuk penelitian sejenis di waktu kedepan untuk bisa menggunakan dataset yang lebih banyak untuk penarikan kesimpulan terhadap hasil penghitungan dengan algoritme K-Means secara lebih aktual.

Daftar Referensi

- [1] L. Izzah and A. Jananto, "Penerapan Algoritma K-Means Clustering Untuk Perencanaan Kebutuhan Obat Di Klinik Citra Medika," *Progresif J. Ilm. Komput.*, vol. 18, no. 1, p. 69, 2022, doi: 10.35889/progresif.v18i1.769.
- [2] R. Aprillina, A. Nurlifa, A. Haryoko, R. E. Putri, and A. N. Rosalita, "Clustering Daerah Rawan Banjir Di Kabupaten Tuban Dengan K-Means Disertai Visualisasi," vol. 1, no. 1, pp. 40–51, 2022.
- [3] Z. I. Alfianti, "Pengelompokan Wilayah Penyebaran Covid-19 Di Kabupaten Karawang Menggunakan Algoritma K-Means," *J. Ilm. Inform. Komput.*, vol. 26, no. 2, pp. 111–122,

- 2021, doi: 10.35760/ik.2021.v26i2.4155.
- [4] N. Dwitri *et al.*, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia," *J. Teknol. Inf.*, vol. 4, no. 1, 2020.
- [5] D. Gultom, I. Gunawan, I. Purnamasari, S. R. Andani, and Z. A. Siregar, "Penerapan Algoritma K-Means Dalam Pengelompokan Kepadatan Penduduk Menurut Kecamatan di Kabupaten Simalungun," *TIN Terap. Inform. Nusant.*, vol. 2, no. 10, pp. 622–628, 2022, doi: 10.47065/tin.v2i10.1375.
- [6] L. Y. Hutabarat, I. Gunawan, I. Purnamasari, M. Safii, and W. Saputra, "Penerapan Algoritma K-Means Dalam Pengelompokan Jumlah Penduduk Berdasarkan Kelurahan Di Kota Pematangsiantar," *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 20–26, 2022, doi: 10.35960/ikomti.v2i2.704.
- [7] S. Wulandari, "Clustering Kecamatan Di Kota Bandung Berdasarkan Indikator Jumlah Penduduk Dengan Menggunakan Algoritma K-Means," *Semin. Nas. Ris. dan Teknol. (SEMNAS RISTEK)*, pp. 128–132, 2020.
- [8] S. S. Informasi and S. T. Dharma, "Data mining Pengelompokan Jumlah Migrasi Penduduk Pertahun dengan Menggunakan Algoritma K-Means Clustering Desa Juhar Kecamatan Juhar " Implementasi," vol. 3, no. 1, 2020.
- [9] P. Anggota *et al.*, "Penerapan K-Means Clustering Pada Penerapan K-Means Clustering Pada Pendaftaran Anggota Perpustakaan Di Dinas Kearsipan Dan Perpustakaan Provinsi Jawa Tengah," *Din. Inform.*, vol. 13, no. 1, 2021.
- [10] A. Asroni, H. Fitri, and E. Prasetyo, "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)," *Semesta Tek.*, vol. 21, no. 1, 2018, doi: 10.18196/st.211211.
- [11] C. S. D. B. Sembiring, L. Hanum, and S. P. Tamba, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri)," *J. Sist. Inf. dan Ilmu Komput. Prima (JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 80–85, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2393.
- [12] Nisar, Wasilah, and H. Kusumajaya, "Pemanfaatan K Means Clustering dalam Pengelompokan Judul Skripsi," *J. JUPITER*, vol. 14, no. 1, pp. 19–26, 2022.
- [13] I. Nasution, A. Perdana Windarto, and M. Fauzan, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi," *Technol. Sci. (BITS)*, vol. 2, no. 2, pp. 76–83, 2020, [Online]. Available: <https://www.bps.go.id>.
- [14] N. A. Rahmalinda and A. Jananto, "Penerapan Metode K-Means Clustering Dalam Menentukan Strategi Promosi Berdasarkan Data Penerimaan Mahasiswa Baru," *J. Tekno Kompak*, vol. 16, no. 2, pp. 163–175, 2022.
- [15] P. Marpaung and R. F. Siahaan, "Penerapan Algoritma K-Means Clustering Untuk Pemetaan Kepadatan Penduduk Berdasarkan Jumlah Penduduk Kota Medan," *J. Sains Komput. Inform.*, vol. 5, no. 1, pp. 503–521, 2021.
- [16] S. Mutdilah, I. Cholissodin, and F. A. Bachtiar, "Improved k-Means (Studi Kasus : Kabupaten Malang)," vol. 4, no. 6, pp. 1850–1857, 2020.
- [17] D. Fitriani, T. N. Padilah, and B. N. Sari, "Penerapan Algoritma K-Means Dalam Pengelompokan Kesejahteraan Rakyat Berdasarkan Kecamatan di Kabupaten Karawang," *Progresif J. Ilm. Komput.*, vol. 17, no. 2, p. 73, 2021, doi: 10.35889/progresif.v17i2.649.
- [18] F. H. Mardiansjah, W. Handayani, and J. S. Setyono, "Pertumbuhan Penduduk Perkotaan dan Perkembangan Pola Distribusinya pada Kawasan Metropolitan Surakarta," *J. Wil. dan Lingkung.*, vol. 6, no. 3, p. 215, 2018, doi: 10.14710/jwl.6.3.215-233.