# OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN METODE ELBOW UNTUK PENGELOMPOKAN RUMAH TANGGA MISKIN DI SURABAYA

### **SKRIPSI**



Disusun Oleh
FITRIANINGSIH
H72216030

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA

2021

## PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : FTTRIANINGSIH

NIM : H72216030

Program Studi : Matematika

Angkatan : 2016

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul " OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN METODE ELBOW UNTUK PENGELOMPOKAN RUMAH TANGGA MISKIN DI SURABAYA". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 14 Juni 2021

Yang menyatakan,

FITRIANINGSIH NIM. H72216030

# LEMBAR PERSETUJUAN PEMBIMBING

## Skripsi oleh

Nama : FITRIANINGSIH

NIM : H72216030

Judul Skripsi : OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN

METODE ELBOW UNTUK PENGELOMPOKAN RU-

MAH TANGGA MISKIN DI SURABAYA

telah diperiksa dan disetujui untuk diujikan.

Surabaya, 14 Juni 2021

Pembimbing

ARIS FANANI M.Kom

NIP. 198701272014031002

# PENGESAHAN TIM PENGUJI SKRIPSI

#### Skripsi oleh

Nama : FITRIANINGSIH

NIM : H72216030

Judul Skripsi : OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN

METODE ELBOW UNTUK PENGELOMPOKAN RU-

MAH TANGGA MISKIN DI SURABAYA

Telah dipertahankan di depan Tim Penguji pada tanggal 1 Juli 2021

Mengesahkan, Tim Penguji

Penguji I

ARIS FANANI M.Kom

NIP. 198701272014031002

Penguji I

Nurissaidah Ulinnuha, M.Kom

NIP. 199011022014032004

Penguji III

Wika Dianita Utami, M.Sc

NIP. 199206102018012003

Penguji I)

Dr. Abdullo Hamid, M.Pd

NIP. 198508282014031003

Mengetahui,

Dekan Fakultas Sains dan Teknologi

IN Sunan Ampel Surabaya

r Hj. Evi Fatimatur Rusyidah, M.Ag

MP. 197312272005012003



# **KEMENTERIAN AGAMA** UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA **PERPUSTAKAAN**

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300 E-Mail: perpus@uinsby.ac.id

#### LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:						
Nama	: Fitrianingsih					
NIM : H72216030 Fakultas/Jurusan : Sains dan Teknologi / Matematika						
			E-mail address : fitrianingsihfitri53@gmail.com			
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :  ✓ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain ()  yang berjudul :  OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN METODE ELBOW UNTUK						
PENGELOMPOK	KAN RUMAH TANGGA MISKIN DI SURABAYA					
Perpustakaan UIN mengelolanya da menampilkan/men akademis tanpa pe	yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, lam bentuk pangkalan data (database), mendistribusikannya, dan npublikasikannya di Internet atau media lain secara <i>fulltext</i> untuk kepentingan erlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai an atau penerbit yang bersangkutan.					
	uk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN baya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta saya ini.					

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 10 Agustus 2021

Penulis

#### **ABSTRAK**

# OPTIMASI K-MEANS CLUSTERING MENGGUNAKAN METODE ELBOW UNTUK PENGELOMPOKAN RUMAH TANGGA MISKIN DI SURABAYA

Kemiskinan merupakan salah satu permasalahan serius yang dihadapi setiap negara, termasuk Indonesia. Terlebih pada saat pandemi COVID-19 yang dialami oleh Indonesia saat ini, semakin memperparah angka kemiskinan. Berbagai program kebijakan saat ini telah diambil pemerintah sebagai bentuk upaya dalam pengentasan kemiskinan, termasuk pemberian bantuan sosial kepada masyarakat miskin. Akan tetapi, bantuan sosial yang diberikan sering kali tidak tepat sasaran atau tidak sesuai dengan yang dibutuhkan masayarakat. Hal tersebut dapat mempengaruhi program pengentasan kemiskinan tidak efisien dan berjalan lebih lamban. Penelitian dilakukan dengan tujuan mengelompokkan masyarakat miskin sesuai dengan kategori, menggunakan metode K-Means yang dioptimasi menggunakan metode Elbow. Pengukuran jarak menggunakan metode Eucledian. Pada setiap kategori diperoleh hasil tingkat kemiskinan paling tinggi terdapat pada kategori bangunan dengan jumlah rumah tangga sangat miskin sebanyak 656, sedangkan tingkat kemiskinan paling rendah terdapat pada kategori pangan dengan jumlah rumah tangga miskin sebanyak 38. Hasil validasi berdasarkan silhouette index yang lebih dari 0.5 dan davies-bouldin index yang mendekati 0 pada setiap kategori menunjukkan cluster yang terbetuk dalam penelitian ini layak digunakan.

Kata kunci: K-Means, Elbow, Kemiskinan, Cluster, Eucledian

#### **ABSTRACT**

# OPTIMIZING K-MEANS CLUSTERING USING ELBOW METHOD FOR CLASSIFICATION OF POOR HOUSEHOLDS IN SURABAYA

Poverty is one of the main problems faced by every country, including Indonesia. What happened during the COVID-19 pandemic experienced by Indonesia at this time, further exacerbated the poverty rate. Various policy programs have been taken by the government as a form of poverty alleviation, including the provision of social assistance to the poor. However, the social assistance provided is often not well targeted or not in accordance with the needs of the community. This can affect poverty alleviation programs to be inefficient and run more slowly. The research was conducted with the aim of classifying the poor according to categories, using the K-Means method which was optimized using the *Elbow* method. Measure the distance using the *Eucledian* method. In each category, the highest poverty rate was found in the building category with 656 very poor households, while the lowest poverty level was found in the food category with 38 poor households. The validation results were based on the *silhouette index* which more than 0.5 and *davies-bouldin index* which is close to 0 in each category indicates that the *cluster* formed in this study is feasible to use.

Keywords: K-Means, Elbow, Poverty, Cluster, Eucledian

## **DAFTAR ISI**

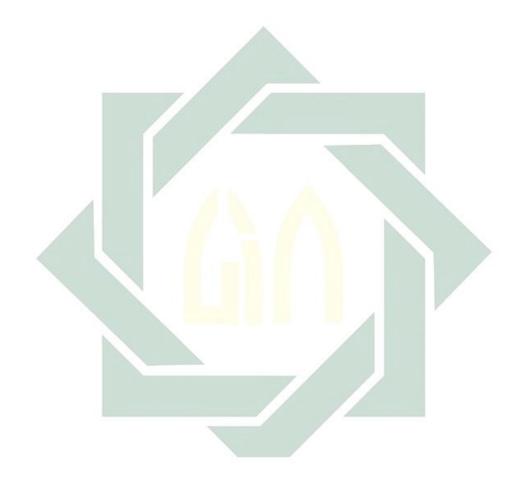
HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI	iii
HALAMAN PERNYATAAN KEASLIAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vi
DAFTAR TABEL	iii
DAFTAR TABEL	X
ABSTRAK	xi
ABSTRACT	xii
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	6
1.3. Tujuan Penelitian	6
1.4. Manfaat Penelitian	7
1.5. Batasan Masalah	7
1.6. Sistematika Penulisan	8
II TINJAUAN PUSTAKA	10
2.1. Kemiskinan	10
2.2. Kemiskinan dalam Perspektif Islam	12
2.3. Analisis Clustering	16
2.4. K-Means	16
2.5. Metode <i>Elbow</i>	18
2.6. Validitas <i>Cluster</i>	19
2.7. Davies-Bouldin Index (DBI)	19
2.8. Silhouette Index	21
HI METODE PENELITIAN	24

	3.1.	Jenis Penelitian	24	
	3.2.	Metode Pengumpulan Data	24	
	3.3.	. Variabel Penelitian		
	3.4.	. Tahap Penelitian		
IV	HAS	SIL DAN PEMBAHASAN	30	
	4.1.	Deskripsi Data	30	
	4.2.	2. Menentukan Jumlah K Cluster Menggunakan Metode Elbow 30		
	4.3.	Perhitungan Analisis <i>Cluster</i> Menggunakan Metode K-Means	36	
	4.4.	Perhitungan Evaluasi Cluster	42	
	4.5.	Interpretasi Cluster	48	
		4.5.1. Kategori Kondisi Bangunan	49	
		4.5.2. Kategori Pangan	51	
		4.5.3. Kategori Kesehatan	52	
		4.5.4. Kategori Karakteristik Kepala Rumah Tangga	53	
		4.5.5. Kategori Kepemilikan Aset	55	
	4.6.	8	57	
$\mathbf{V}$	PEN	UTUP	<b>59</b>	
	5.1.	Simpulan	59	
	5.2.	Saran	60	
DA	FTA	R PUSTAKA	61	
A	SKR	IP PROGRAM PYTHON	64	

## **DAFTAR TABEL**

2.1	nilai ukur Silhouette Index	23
3.1	Variabel Penelitian	25
4.1	WSS Pada Kategori Kondisi Bangunan	31
4.2	WSS Pada Kategori Pangan	32
4.3	WSS Pada Kategori Kesehatan	33
4.4	WSS Pada Kategori Karakteritik Kepala Rumah Tangga	
4.5	WSS Pada Kategori Kepemilikan Aset	35
4.6	Data Kepemilikan Aset	37
4.7	Centroid Pada Iterasi ke-1	37
4.8	Hasil Perhitungan Jarak Pada Iterasi ke-1	39
4.9	Rata-Rata <i>Cluster</i> ke-1 Pada Iterasi ke-1	39
4.10	Rata-Rata <i>Cluster</i> ke-2 Pada Iterasi ke-1	40
4.11	Rata-Rata <i>Cluster</i> ke-3 Pada Iterasi ke-1	40
4.12	Rata-Rata <i>Cluster</i> ke-4 Pada Iterasi ke-1	40
4.13	Centroid Pada Iterasi ke-1	41
4.14	Hasil Perhitungan Jarak Pada Iterasi ke-2	41
4.15	Hasil SSW Pada Setiap <i>Cluster</i>	43
4.16	Hasil SSB Antar Centroid	44
4.17	Hasil Evaluasi Cluster Pada Setiap Kategori	48
4.18	Hasil Centroid Cluster Pada Kategori Kondisi Bangunan	49
4.19	Hasil Cluster Pada Kategori Kondisi Bangunan	50
4.20	Hasil Centroid Pada Kategori Pangan	51
4.21	Hasil Cluster Pada Kategori Pangan	52
4.22	Hasil Centroid Pada Kategori Kesehatan	52
4.23	Hasil Cluster Pada Kategori Kesehatan	53
4.24	Hasil Centroid Pada Kategori Karakterisritik Kepala Rumah Tagga .	54

4.25	Hasil Centroid Pada Kategori Karakteristik Kepala Rumah Tangga .	55
4.26	Hasil Centroid Pada Kategori Kepemilikan Aset	56
4.27	Hasil Centroid Pada Kategori Kepemilikan Aset	56



# DAFTAR GAMBAR

3.1	Flawchart K-Means	28
4.1	Grafik WSS Pada Kategori Kondisi Bangunan	32
4.2	Grafik WSS Pada Kategori Pangan	33
4.3	Grafik WSS Pada Kategori Kesehatan	34
4.4	Grafik WSS Pada Kategori Karakteritik Kepala Rumah Tangga	35
15	Grafik WSS Dada Katagori Kanamilikan Asat	26

#### **BABI**

#### **PENDAHULUAN**

#### 1.1. Latar Belakang Masalah

Kemiskinan merupakan salah satu permasalahan serius yang dihadapi setiap negara, termasuk Indonesia. Berdasarkan data Badan Pusat Statistik, persentase jumlah penduduk miskin di Indonesia pada tahun 2015-2019 terus mengalami penurunan, dari 4,31% menjadi 3,81%. Akan tetapi, pada September 2019 sampai Maret 2020 persentase meningkat sebesar 0,56% yaitu dengan peningkatan jumlah penduduk miskin dari 1,28 juta orang menjadi 1,63 juta orang. Jika dibedakan antara penduduk miskin di daerah dan di perkotaan, maka peningkatan persentase penduduk miskin di perkotaan lebih besar. Pada September 2019 – Maret 2020 penduduk miskin di perkotaan mengalami peningkatan sebesar 0,82% (dari 6,56% menjadi 7,38%), sedangkan penduduk miskin di daerah mengalami peningkatan sebesar 0,22% (dari 12,60% menjadi 12,82%) (Badan Pusat Statistik, 2020). Salah satu penyebab kemiskinan meningkat pada tahun 2020 yaitu kebijakan yang diambil untuk mencegah penyebaran Covid-19, diantara kebijakan tersebut yaitu Pembatasan Sosial Bersekala Besar (PSBB) yang mengakibatkan perekonomian di Indonesia menurun. Kebijakan pemerintah tersebut juga berdampak pada jumlah pemutusan hubungan kerja semakin meningkat hingga 3,5 juta orang lebih, dikarenakan sebagian peruasahaan tidak dapat melakukan kegiatan produksi (CNBC Indonesia, 2020).

Dalam menghadapi masalah kemiskinan yang semakin meningkat di Indo-

nesia, Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K) telah melakukan penetapan Rencana Pembangunan Jangka Menengah Nasional (RPJM) 2020-2024. TNP2K merupakan lembaga yang dibentuk khusus oleh presiden dan diketuai langsung oleh wakil presiden RI untuk meningkatkan penanganan program pengentasan kemiskinan di Indonesia. Adapun target pengurangan kemiskinan yang ditetapkan oleh TNP2K dalam RPJM mencapai 7% target moderat dan 6,5% target optimis pada tahun 2024. Jika penduduk miskin pada 2024 diperkirakan sebanyak 19,75 juta orang, maka dibutuhkan penurunan penduduk miskin sebesar 5,04 juta sampai 6,45 juta orang untuk memenuhi target. Bentuk upaya yang dilakukan oleh TNP2K diantanya meningkatkan pendapatan dan menurunkan beban pengeluaran masyakat melalui berbagai perbaikan program bantuan sosial seperti Pogram Indonesia Pintar Kuliah (PIP-K), Program Keluarga Harapan (PKH), Program Bantuan Pangan (Rastra dan Bantuan Pangan Non-Tunai), Jaminan Kesehatan Naional (JKN), bantuan subsidi listrik, LPG dan masih banyak bantuan yang diberikan pemerintah (Tim Nasional Percepatan Pengentasan Kemiskinan, 2020).

Berbagai program bantuan yang disediakan pemerintah diberikan kepada rumah tangga yang tergolong kurang mampu, untuk membantu meringankan beban biaya pengeluaran rumah tangga tersebut. Akan tetapi, bantuan yang dibutuhkan setiap rumah tangga yang tergolong tidak mampu memiliki proporsi yang berbedabeda. Dengan mempertajam ketepatan sasaran dalam menyalurkan berbagai jenis bantuan yang telah direncanakan, dapat membantu pemerintah dalam mencapai target angka pengentasan kemiskinan serta membantu masyarakat kurang mampu mendapatkan hak-hak yang seharusnya mereka peroleh. Allah SWT berfirman dalam Al-Quran Surat An-Nisa ayat 58:

Terjemah: "Sesungguhnya Allah menyuruh kamu menyampaikan amanat kepada yang berhak menerimanya, dan (menyuruh kamu) apabila menetapkan hukum di antara manusia supaya kamu menetapkan dengan adil. Sesungguhnya Allah memberi pengajaran yang sebaik-baiknya kepadamu. Sesungguhnya Allah adalah Maha Mendengar lagi Maha Melihat."

Pada ayat tersebut menjelaskan, pentingnya menjalankan amanat secara cepat dan tepat kepada pemiliknya. Jika ditinjau dari penulisan pada awal ayat yang terdapat penyebutan nama Allah SWT secara langsung dalam memerintah, hal tersebut dapat diartikan bahwa kalimat perintah di ayat ini sangat ditekankan. Dalam tafsir Misbah dijelaskan mengenai pengertian "amanat yaitu sesuatu yang diserahkan kepada pihak lain untuk dipelihara dan dikembalikan bila tiba saatnya atau bila diminta oleh pemiliknya" (Shihab, 2020). Berdasarkan pengertian tersebut, pemerintah selaku pengemban amanat rakyat berperan penting dalam menyampaikan amanat dengan semaksimal mungkin. Masayarakat berhak menerima bantuan dari pemerintah sesuai dengan proporsi masing-masing. Akan tetapi, selama ini masih banyak pemberian bantuan yang masih kurang tepat sasaran bahkan sampai terjadi tumpang tindih dalam penerimaan bantuan.

Berbagai masalah kemiskinan terjadi di daerah Indonesia, begitu juga dengan kota Surabaya yang merupakam salah satu kota metropolitan di Indonesia. Berdasarasarkan survei yang dilakukan oleh LaporCovid-19 bersama Social Resilience Lab Nanyang Technological University (NTU), yang dilakukan pada 19 Juni – 10 Juli 2020 di Surabaya dengan jumlah responden sebanyak 2.895. Dalam survei tersebut terdapat 68% responden yang menyatakan penghasilannya menurun dras-

tis, 18% penghasilannya sedikit berkurang dan 14% mengatakan tidak berkurang (Lidwina, 2020). Tidak hanya mengenai permasalahan tersebut, sebagian warga Surabaya juga masih merasakan adanya bantuan sosial yang tidak tepat saasaran, seperti adanya warga yang masih menirima bantuan padahal orang tersebut sudah meninggal (Andira, 2020).

Dalam upaya pengentasan kemiskinan sejatinya pemerintah ataupun masyarakat saling memiliki peran penting di dalamnya. Banyak upaya yang dapat dilakukan dalam menghadapi permasalahan kemiskinan, salah satunya yaitu dengan melakukan pengelompokan rumah tangga sesuai dengan kesamaan karakteristik yang dimiliki. Dari pengelompokan rumah tangga tersebut, diharapkan dapat diketahui jenis bantuan apa yang dibutuhkan setiap kelompok. Algoritma K-Means merupakan salah satu metode *clustering* yang dapat digunakan, karena termasuk dalam non-herarki *clustering* sehingga jumlah K *cluster* dapat ditentukan sesuai dengan yang dibutuhkan. Akan tetapi dalam metode K-Means masih terdapat kelemahan yaitu dalam menganalisa jumlah K *cluster* yang baik digunakan. Oleh sebab itu, perlu adanya metode untuk menentukan jumlah K *cluster* yang baik. Salah satu metode yang dapat digunakan yaitu metode *elbow*, kemampuannya dalam menganalisa jumlah K *cluster* dapat digunakan untuk mengoptimalkan kinerja K-Means.

Beberapa penelitian yang pernah dilakukan sebelumnya mengenai klasterisasi rumah tangga miskin dilakukan oleh Dwi Rahayu Utami (2018) dengan judul "Aplikasi Monitoring Keluarga Miskin Menggunakan Metode K-Means *Clustering* Berbasis Mobile GIS (Studi Kasus: PKH Kecamatan Kedungkandang Kota Malang)", dalam penelitian tersebut menjelaskan mengenai pembuatan aplikasi mobile dengan menggunakan algoritma K-Means *Clustering* untuk mengelompokan keluarga miskin. Adapun nilai akurasi yang dihasilkan dari pengelompokan sebesar

92,8% atau bisa disebut dengan Excelent classification (Utami, 2018).

Penelitian juga dilakukan oleh Nugroho Irawan Febianto dan Nico Dias Palasara (2019) dengan judul penelitian "Analisis *Clustering* K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018". Pada kesimpulan penelitian tersebut menghasilkan 5 *cluster* yang memetakan wilayah provinsi Jawa Barat berdasar nilai rata-rata tertinggi dan terendah yang diperoleh dari indikator kemiskinan (Febianto dan Nico, 2019).

Carina Kallestal, dkk (2020) melakukan penelitian dengan judul "Assessing the Multiple Dimensions of Poverty. Data Mining Approaches to the 2004-14 Health and Demographic Surveillance System in Cuatro Santos Nicaragua". Algoritma yang digunakan pada penelitian tersebut yaitu K-Means *cluster*, karena dinilai lebih mudah digunakan dan ditafsirkan. Variabel pada penelitian tersebut dibagi menjadi 5 kelompok, yaitu Kebutuhan dasar yang tidak terpenuhi, aset, kerawanan pangan, intervensi dan variabel individu turunan. Penelitian tersebut menghasilkan 6 klaster yang menunjukkan adanya kemiskinan multidimensi. Hasil penelitian tersebut telah dikonfirmasi ketepatannya melalui diskusi yang dilakukan peneliti bersama tokoh masyarakat setempat termasuk tokoh kesehatan, keamanan serta perwakilan dari masyarakat awam (Kallestal et al., 2020).

Penelitian mengenai validitas *cluster* yang dilakukan oleh Khairatia dkk (2019) dengan judul "Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan Kmeans MMCA" yang membandingkan indeks validitas *silhouette index*, *Davies Bouldin Index*, *Dunn* dan *Calinski-Harabasz* menyatakan *silhouette index*, *Davies Bouldin Index* dan *Calinski-Harabasz* memiliki hasil yang lebih baik dibandingkan *Dunn*. Hasil tersebut didapat dari algoritma *clustering* K-Means, K-Means Enhanced dan K-Means MMCA sebagai pembanding hasil validitas *cluster* (Khai-

rati dkk., 2019).

Penelitian juga dilakukan oleh Nisa Hanum dkk (2020) dengan judul peneltian "Segmentasi pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python" dalam penelitian tersebut menjelaskan mengenai pemilihan jumlah K *cluster* terbaik menggunkan metode *elbow*. Adapun metode yang digunakan untuk evaluasi *cluster* yaitu *Silhouette Index* dan *Davies Bouldin Index* dan *Calinski-Harabash*. Penelitian tesebut menghasilkan jumlah K *cluster* terbaik yaitu 3, dengan nilai *Silhouette Index* mendekati 1 dan nilai *Davies Bouldin Index* mendekati 0 yang berarti hasil dari *cluster* tersebut termasuk dalam klaster yang kuat (Harani, dkk., 2020).

Berdasarkan penjelasan di atas, penulis tertarik melakukan penelitian yang berjudul "Optimasi *K-Means Clustering* Menggunakan Metode *Elbow* untuk Pengelompokan Rumah Tangga Miskin di Surabaya"

#### 1.2. Rumusan Masalah

Berdasarkan uraian dari latar belakang yang telah dijelaskan diatas maka rumusan masalah dari penelitian ini antara lain:

- 1. Bagaimana hasil optimasi metode K-Means *clustering* menggunakan metode *elbow* untuk mengelompokkan rumah tangga miskin agar bantuan sesuai dengan tujuan?
- 2. Bagaimana validasi pengelompokan rumah tangga miskin di Surabaya menggunakan analisis *silhouette index* dan *davies-bouldin index*?

#### 1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini antara lain:

- Mengetahui hasil optimasi metode K-Means clustering menggunakan metode elbow untuk. mengelompokkan rumah tangga miskin agar sesuai dengan tujuan
- 2. Mengetahui validasi pengelompokan rumah tangga miskin di Surabaya menggunakan analisis *silhouette index* dan *davies-bouldin index*.

#### 1.4. Manfaat Penelitian

- 1. Bagi Penulis: Penelitian ini dapat menambah pengetahuan bagi penulis mengenai penerapan metode K-Means *clustering* dalam kehidupan bersosial, salah satunya yaitu dalam pengelompokan rumah tangga miskin menggunakan variabel tertentu.
- 2. Bagi Pembaca: Penlitian ini dapat dijadikan bahan referensi mengenai penerapan analisis K-Means *clustering* dalam pengelompokan rumah tangga miskin menggunakan variabel tertentu.
- 3. Bagi Pemerintah Kota: Penelitian ini dapat memberikan gambaran mengenai keadaan rumah tangga miskin di Surabaya, sehingga pemerintah kota dapat memberikan bantuan yang sesuai dengan kebutuhan setiap rumah tangga yang tergolong tidak mampu.

#### 1.5. Batasan Masalah

Dalam penelitian ini, peneliti membatasi masalah sebagai berikut:

1. Data yang digunakan yaitu data SUSENAS di Kota Surabaya tahun 2018 yang meliputi luas bangunan  $(m^2)$ , jenis lantai, jenis dinding, fasilitas buang air besar, melewatkan makan, makan lebih sedikit, kehabisan makanan, memiliki

BPJS PBI, memiliki BPJS Non PBI, memiliki jamkesmas, memiliki asuransi kesehatan swasta, memiliki asuransi kesehatan perusahaan, tidak punya asuransi kesehatan, pendidikan terakhir kepala rumah tangga, lapangan pekerjaan kepala rumah tangga, memiliki kulkas, memiliki AC, memiliki mobil.

2. Pengukuran jarak menggunakan *Euclidean*. Algoritma yang digunakan yaitu K-Means *clustering* serta validasi pengelompokan menggunakan analisis *silhouette index* dan *davies-bouldin index*.

#### 1.6. Sistematika Penulisan

Sistematika penulisan yang di gunakan dalam menulis skripsi ini sebagai berikut:

#### 1. BAB I: PENDAHULUAN

Bab ini menjelaskan t<mark>entang latar</mark> b<mark>ela</mark>kang <mark>pe</mark>mbuatan skripsi, rumusan masalah, tujuan penelitian, batasan masalah dan sistematika penulisan skripsi.

#### 2. BAB II: TINJAUAN PUSTAKA

Bab ini berisi penjelasan teori tentang sistem yang digunakan dalam penelitian ialah konsep kemiskinan, analisis *cluster*, metode K-Means *cluster*, uji validasi *cluster* menggunakan analisis *silhouette index* dan *davies-bouldin inde*.

#### 3. BAB III: METODE PENELITIAN

Bab ini memuat uraian tentang jenis penelitian, teknik pengumpulan data, variabel penlitian, serta tahap penelitian.

#### 4. BAB IV: HASIL DAN PEMBAHASAN

Bab ini menjelaskan deskripsi mengenai hasil pengklasteran rumah tangga

miskin di Surabaya menggunakan metode K-Means yang dioptimasi dengan metode *elbow*.

#### 5. BAB V: SIMPULAN

Bab ini menjelaskan secara ringkas mengenai hasil yang didapat, serta saran bagi penelitian selanjutnya.



#### **BABII**

#### TINJAUAN PUSTAKA

#### 2.1. Kemiskinan

Kemiskinan adalah suatu keadaan dimana kebutuhan dasar hidup tidak dapat dipenuhi secara layak. Adapun keterbatasan yang dimaksut yaitu akses informasi, pendapatan, kondisi kesehatan, keterampilan, ekonomi, penguasaan aset, serta pengukuran yang bersifat ekonomi. Pengukuran tersebut dilakukan dengan cara melakukan pendekatan moneter. Badan Pusat Statistik menggunakan data mengenai pengeluaran rumah tangga untuk melakukan pendekatan pendapatan dalam rumah tangga. Dari data pengeluaran rumah tangga tersebut, kemudian dibandingkan dengan batas nilai tukar rupiah yang dibutuhkan dalam memenuhi minimum kebutuhan hidup batas yang telah ditentukan disebut dengan garis kemiskinan. Penduduk masuk dalam kategori miskin jika pengeluaran berada pada bawah garis kemiskinan yang telah ditentukan. Adapun garis kemiskinan yang ditentukan pada setiap daerah berbeda-beda (Isdijoso dkk., 2016).

Pada tahun 2005 BPS melakukan pendataan yang menghasilkan daftar nama dan alamat rumah tangga miskin. Data rumah tangga miskin tersebut dapat digolongkan berdasarkan tingkat keparahannya, yaitu rumah tangga sangat miskin, miskin dan hampir miskin. Untuk menentukan rumah tangga miskin BPS menggunakan 14 variabel yang telah dimuttakhirkan pada Oktober 2008, sebagai berikut (Isdijoso dkk., 2016):

#### 1. Luas lantai per kapita

- 2. Jenis lantai
- 3. Jenis dinding
- 4. Fasilitas buang air besar
- 5. Sumber air minum
- 6. Sumber penerangan
- 7. Jenis bahan bakar
- 8. Pembelian daging/ayam/susu
- 9. Frekuensi makan
- 10. Pembelian pakaian baru
- 11. Kemampuan berobat
- 12. Lapangan usaha kepala rumah tangga
- 13. Pendidikan terakhir kepala rumah tangga
- 14. Aset yang dimiliki

Menurut pendataan sosial ekonomi tahun 2005, rumah tangga termasuk dalam kategori miskin apabila (Isdijoso dkk., 2016):

- 1. Luas lantai bangunan tempat tempat tinggalnya kurang dari 8  $m^2$  per orang
- 2. Lantai bangunan tempat tinggalnya terbuat dari tanah/bambu/kayu murahan
- 3. Dinding bangunan tempat tinggalnya terbuat dari bambu/rumbia/kayu berkualitas rendah atau tembok tanpa di plaster

- 4. Tidak memiliki fasilitas buang air besa/bersama-sama rumah tangga lain menggunakan satu jamban
- 5. Sumber penerangan rumah tangga tidak menggunakan listrik
- 6. Air minum berasal dari sumur/mata air yang tidak terlindung/sungai/air hujan
- 7. Bahan bakar untuk memasak sehari-hari adalah kayu bakar/arang/minyak tanah
- 8. Hanya mengonsumsi daging/susu/ayam satu kali dalam seminggu
- 9. Hanya membeli satu stel pakaian baru dalam setahun
- 10. Hanya mampu makan satu/dua kali dalam sehari
- 11. Tidak sanggup membayar biaya pengobatan di puskesmas/politeknik
- 12. Sumber penghasilan kepala rumah tangga adalah: petani dengan luas lahan 0,5 ha, buruh tani, nelayan, buruh bangunan, buruh perkebunan, atau pekerjaan lainnyadengan pendapatan di bawah Rp 600.000 per bulan
- Pendidikan terakhir kepala rumah tangga: tidak sekolah/ tidak tamat sekolah dasar (SD)/hanya SD
- 14. Tidak memiliki tabungan/barang yang mudah dijual dengan nilai minimal Rp 500.000 seperti sepeda motor (kredit/nonkredit), emas, hewan ternak, kapal motor ataupun barang modal lainnya.

#### 2.2. Kemiskinan dalam Perspektif Islam

Dalam Al-Quran kata miskin juga dikatakan sebagai "Maskanat" (Kemiskinan) yang merupakan *mashdar* dari *fi'il madhi* "Sakana". Menurut Al-Raghib

Al-Ashfahani kata miskin juga berkaitan dengan kata "Al-Maskanaat" yang merupakan sifat *musyabbihat* (sesuatu yang tidak terikat oleh 3 zaman / waktu yakni masa lampau, sekarang dan akan datang) dengan arti suatu keadaan manusia yang tidak memiliki harta benda. Kemiskinan di dalam Al-Quran biasa berhubungan dengan persoalan manusia, dalam hal ini mengenai kemiskinan ekonomi atau materi. Adapun fakir miskin menurut islam ialah suatu keadaan sesorang yang tidak memiliki sumber mata pencaharian atau seseorang yang memiliki sumber mata pencaharian akan tetapi tidak mampu memenuhi kebutuhan hidupnya dengan layak. Jika ditinjau dalam Bahasa Arab, seseorang yang tergolong miskin berarti sesorang tersebut berada dalam keadaan kefakiran yang sangat (Ulya, 2018). Dalam Al-Quran dijelaskan mengenai kemiskinan dalam surat Al-Baqarah ayat 273:

Artinya: "(Apa yang kamu infakkan) adalah untuk orang-orang fakir yang terhalang (usahanya karena jihad) di jalan Allah, sehingga dia yang tidak dapat berusaha di bumi; (orang lain) yang tidak tahu, menyangka bahwa mereka adalah orang-orang kaya karena mereka menjaga diri (dari meminta-minta). Engkau (Muhammad) mengenal mereka dari ciri-cirinya, mereka tidak meminta secara paksa kepada orang lain. Apa pun harta yang baik yang kamu infakkan, sungguh, Allah Maha Mengetahui."

Ayat tersebut berkaitan dengan Hadist Rasulullah SAW yang diriwayatkan dalam hadist Muslim 1772, sebagai berikut:

حَدَّتَنَا قُتَيْبَةُ بْنُ سَعِيدِ حَدَّتَنَا الْمُغِيرَةُ يَعْنِي الْحِزَامِيَّ عَنْ أَبِي الرِّنَادِ عَنْ الْأَعْرَجِ عَنْ أَبِي هُرَيْرَةَ أَنَّ رَسُولَ اللهِ صَلِّى اللَّاسِ فَتَرُدُّهُ اللَّقُمَةُ وَاللَّهُمَتَانِ اللهِ صَلِّى اللهِ عَلَى اللهِ عَلَى النَّاسِ فَتَرُدُّهُ اللَّقُمَةُ وَاللَّهُمَتَانِ وَالنَّمْرَةُ وَالنَّمْرَةُ وَالنَّمْرَتَانِ قَالُوا فَمَا الْمِسْكِينُ يَا رَسُولَ اللهِ قَالَ الّذِي لَا يَجِدُ غِنَى يُغْنِيهِ وَلَا يُفْطَنُ لَهُ فَيُتَصَدَّقَ عَلَيْهِ وَلَا يَفْطَنُ لَهُ فَيُتَصَدَّقَ عَلَيْهِ وَلَا يَسْأَلُ النَّاسَ شَيْئً

Artinya: Telah menceritakan kepada kami Qutaibah bin Sa'id telah menceritakan kepada kami Al Mughirah Al Hizami dari Abu Zinad dari Al A'raj dari Abu Hurairah bahwa Rasulullah shallallahu 'alaihi wasallam bersabda: "Orang miskin bukanlah mereka yang berkeliling meminta-minta kepada orang banyak, lalu peminta itu diberi sesuap dua suap, atau sebutir dua butir kurma." Para sahabat bertanya, "Kalau begitu, seperti apakah orang yang miskin itu?" Beliau menjawab: "Orang miskin sesungguhnya ialah mereka yang tidak memiliki apa-apa untuk menutupi kebutuhannya, namun keadaannya itu tidak diketahui orang supaya orang bersedekah padanya, dan tidak pula meminta-minta ke sana ke mari."

Ayat Al-Quran dan Hadist tersebut menjelaskan bahwasannya faktor yang melatarbelakangi kemiskinan bukan berasal dari mentalitas seseorang. Hal tersebut dapat diketahui dari Nabi SAW yang menganjurkan untuk memberi sebagian harta kepada orang yang tidak meminta-minta serta senantiasa mencari mata pencaharian dengan sungguh-sungguh, meskipun belum cukup untuk memenuhi kebutuhan hidupnya (Duraesa, 2016).

Dalam Al-Quran banyak yang menjelaskan tentang betapa pentingnya berbagi kepada orang miskin, diantaranya:

Artinya: "Apakah yang memasukkan kamu ke dalam (neraka) Saqar?" Me-

reka menjawab: "Kami dahulu tidak Termasuk orangorang yang mengerjakan shalat, Dan Kami tidak (pula) memberi Makan orang miskin." (Q.S Al-Muddatssir: 42-44)

Artinya: "Tahukah kamu (orang) yang mendustakan agama? Itulah orang yang menghardik anak yatim dan tidak menganjurkan memberi makan orang miskin." (Q.S Al-Maun: 1-3)

Artinya: "Dan apabi<mark>la pada waktu p</mark>emba<mark>gia</mark>n itu hadir beberapa kerabat, anak-anak yatim dan orang-orang miskin, maka berilah mereka dari sebagian harta itu dan ucapkanlah perkataan yang baik kepada mereka." (Q.S An-Nisa: 8)

Ayat tersebut di atas hanya sebagian kecil dari firman Allah SWT yang menegaskan mengenai pentingnya berbagi kepada orang miskin. Permasalahan kemiskinan merupakan suatu permasalahan yang sangat perlu diatasi karena dapat merambah pada permasalahan lain seperti munculnya keterbelakangan baik dalam segi ilmu pengetahuan, teknologi, maupun sosial budaya. Bahkan Ali bin Abi Thalib sampai mengatakan "seandainya kefakiran/kemiskinan itu berwujud manusia, maka niscaya akan kubunuh dia". Kalimat tersebut merupakan kalimat yang menunjukkan bahwasannya agama Islam sangat serius dalam menangani permaslahan kemiskianan bahkan sejak zaman dahulu (Duraesa, 2016).

#### 2.3. Analisis Clustering

Clustering merupakan pengelompokan data/vektor secara partisi berdasarkan karakteristik masing-masing. Data-data hanya masuk pada cluster yang sama jika memiliki karakter yang sama. Clustering juga biasa disebut dengan unsupervised learning karena tidak memiliki ketentuan kelas (Jiawei et al., 2011).

Kelas pada *clustering* tidak memerlukan label ketika pemrosesan, karena label bisa diberikan setelah pengelompokan terbentuk. Oleh sebab itu, *clustering* cocok digunakan pada data yang sulit dicari label kelasnya ketika pembangkitan fitur. Selain itu, *clustering* juga biasa digunakan pada data yang akan dilakukan *supervised* untuk mengakarakteristikkan suatu data. Perlu diketahui bahwa, pendefisian kemiripan (*similarity*) dalam *cluster* harus didasarkan terhadap atribut objek. Algoritma *clustering* yang digunakan tergantung pada proses yang diterapkan pada set data (Prasetyo, 2014).

#### 2.4. K-Means

K-Means merupakan salah satu metode *clustering* non-hierarki yang berfungsi untuk membagi data secara partisi. Pengelompokan data dalam metode K-Means tidak memerlukan kelas, sehingga data-data dikelompokkan langsung sesuai dengan masukan. Adapun masukan yang dimaksut yaitu data dan jumlah K *cluster* yang diinginkan. Data-data yang telah dimasukkan akan dikelompokkan ke dalam K *cluster* yang telah ditentukan. Di dalam setiap *cluster* terdapat *centroid* (titik pusat) yang dipilih secara acak pada tahap pertama. *Centroid* tersebut digunakan untuk mempresentasikan atau menggambarkan *cluster*. Posisi *centroid* dalam *cluster* akan terus berubah-ubah sampai pada posisi optimal (Wanto dkk., 2020). Pada setiap K *cluster* digambarkan oleh titik tunggal dalam r-dimensi ( $R^d$ ). Set yang

merepresentasikan *cluster* atau biasa disebut dengan *centroid*, dinyatakan sebagai berikut (Prasetyo, 2014):

$$\mathcal{C} = \{c_i | j = 1, \dots, K\}$$

Konsep yang digunakan dalam pengelompokan merupakan kesamaan (*similarity*) karakteristik pada data. Akan tetapi, kemiripan pada data-data diukur dengan menggunakan kuantitas ketidakmiripan (*dissimilarity*). Dapat diartikan, data yang memiliki ketidakmiripan (jarak) yang dekat kemungkinan besar akan berada dalam satu *cluster* yang sama. Adapun matriks yang biasa digunakan untuk mengukur ketidakmiripan (jarak) ialah *Euclidean*, dengan persaman sebagai berikut (Prasetyo, 2014):

$$d(x,y) = ||x - y||_2 = \sqrt{\sum_{i=1}^{r} (x_i - y_i)^2}$$
 (2.1)

Dimana:

 $x_i$  = nilai fitur ke-i dari x

 $y_i$  = nilai fitur ke-i dari y

r = jumlah fitur dalam vektor

d = ketidakmiripan

d(x,y) = d(y,x)

Nilai dari  $d_0$  dalam *Euclidean* yaitu 0 atau bisa dikatakan jarak minimal antara dua vektor bernilai 0. Data yang memiliki ketidakmiripan paling kecil terhadap *centroid* akan terpilih kedalam *cluster*. Sedangkan data yang memiliki ketiakmiripan terhadap *centroid* akan direlokasikan ke *centroid* lain yang memiliki jarak lebih

dekat. Relokasi data dalam *cluster* memiliki nilai a anggota sama dengan 1 jika data tersebut menjadi anggota sebuah *cluster* dan bernilai 0 jika tidak masuk dalam anggota *cluster*. Setiap data pada *cluster* hanya memiliki satu nilai yaitu 0 atau 1. Hal tersebut tejadi karena dalam pengelompkan K-Means hanya mengelompokkan data tepat pada satu *cluster* (Prasetyo, 2014).

Untuk relokasi *centroid* terhadap *cluster*, dapat dicari dengan menghitung rata-rata dari keseluruhan data yang terkumpul dalam masing-masing *cluster* dengan persamaan sebagai brikut (Prasetyo, 2014):

$$c_j = \frac{1}{Nk} \sum_{l=1}^{Nk} x_{jl}$$
 (2.2)

Dimana:

Nk = Banyak data yang berada dalam sebuah*cluster* 

 $x_{il}$  = data ke-j dari fitur ke-l

#### 2.5. Metode Elbow

Metode *elbow* merupakan salah satu metode yang dapat digunakan untuk menentukan jumlah *cluster* terbaik, dengan cara melihat persentase setiap *cluster* yang akan membentuk siku pada suatu titik tertentu. Metode *elbow* biasa disajikan dalam bentuk grafik untuk mengetahui lebih jelas siku yang terbentuk. Adapun persamaan yang digunakan dalam metode *elbow* yaitu nilai total WSS (*Within Cluster Sum of Squares*) atau biasa disebut dengan SSE (*Sum Square Error*) dengan rumus sebagai berikut (Purnama, 2019):

$$WSS = \sum_{k=1}^{k} \sum_{x_i \in C_k} (x_i - \varphi_k)^2$$
 (2.3)

#### Keterangan:

 $C_k = K cluster yang terbentuk$ 

k = banyak cluster

 $x_i$  = data x pada fitur ke-i

 $\varphi_k$  = rata-rata K *cluster* pada nilai k (k=1,2,3,...,K)

#### 2.6. Validitas Cluster

Jika dilihat dari analisis *cluster* yang merupakan metode untuk penggalian informasi baru, maka validasi pada *cluster* seolah tak diperlukan. Akan tetapi, jika dilihat dari data yang sama dengan pemilihan jumlah *cluster* yang berbeda maka akan menghasilkan susunan *cluster* yang berbeda pula. Oleh sebab itu, validasi pada *cluster* sangat penting untuk mengevaluaisi *cluster* yang telah terbentuk dengan cara memberi nilai validitas. Adapun cara evaluasi yang digunakan tergantung pada algoritma yang digunakan. Informasi yang dibutuhkan dalam mengevaluasi *cluster* ada dua, yaitu informasi internal (metode *unsupervised*) yang digunakan untuk mengukur seberapa bagus struktur pada *cluster* dan informasi internal (metode *supervised*) yang digunakan untuk mengukur kecocokan struktur *cluster* yang telah didapat dengan struktur eksternal (Tan et al., 2006).

#### 2.7. Davies-Bouldin Index (DBI)

DBI merupakan validitas *cluster* internal yang menggunakan metode berbasis kohesi dan sparasi. Kohesi yang dimaksud adalah jumlah dari kedekatan data dalam *cluster* terhadap *centroidnya*. Nilai DBI didapat dari rata-rata setiap *cluster* 

yang memiliki kesamaan. Kemudian, dilakukan perbandingan nilai DBI yang didapat dari *cluster* yang didapat. Adapun persamaan DBI sebagai berikut (Fhadli 2020):

$$DBI = \frac{1}{K} \sum_{i=1}^{K} \max_{i=j} (R_{i,j})$$
 (2.4)

Dimana:

K = jumlah cluster

 $R_{ij}$  = rasio perbandingan antara ke-i dan ke-j

Nilai  $R_{i,j}$ dapat diketahui melalui persamaan berikut:

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_i, j} \tag{2.5}$$

SSW (Sum of Square Within Cluster) merupakan matriks kohesi yang terdapat dalam cluster ke-i, dengan persamaan sebagai berikut:

$$SSW_i = \frac{1}{m_1} \sum_{j=1}^{m_i} d(x_j, c_i)$$
 (2.6)

Dimana:

 $m_i$  = jumlah data yang berada dalam *cluster* ke-i

 $d(x_j, c_i)$  = jarak data ke-j ke *centroid* yang berada dalam *cluster* ke-i

SSB (Sum of Square Between Cluster) merupakan matriks separasi yang berfungsi untuk mengukur jarak antar cluster, dengan cara mengukur jarak centroid dalam suatu cluster dengan centroid yang berada dalam cluster lain. Berikut persamaan SSB yang digunakan untuk untuk mengukur jarak antara cluster ke-i dan

cluster ke-j:

$$SSB_{i,j} = d(c_i, c_j) \tag{2.7}$$

Dimana:

 $d(c_i, c_j) = \text{jarak } centroid } c_i \text{ ke } centroid } c_j$ 

Cluster dikatakan baik jika memiliki nilai kohesi sekecil mungkin dan nilai separasi sebesar mungkin. Nilai DBI dikatakan baik jika mendekati non-negatif  $\geq$  0. Adapun sifat-sifat yang dimiliki oleh  $R_{ij}$  sebagai berikut (Prasetyo, 2014):

- 1.  $R_{ij} \ge 0$
- $2. R_{ij} = R_{ji}$
- 3. Jika  $SSW_j \geq SSW_r$  dan  $SSB_{i,j} = SSB_{i,r}$ , maka  $R_i, j > R_{i,r}$
- 4. Jika  $SSW_j = SSW_r$  dan  $SSB_{i,j} \leq SSB_{i,r}$ , maka  $R_i, j > R_{i,r}$

#### 2.8. Silhouette Index

Silhouette Index digunakan untuk digunakan untuk memvalidasi data mana yang cocok untuk masuk dalam suatu cluster. Pada umumnya, Silhouette Index didapat dari hasil bagi perhitungan nilai rata-rata selisih separasi dan compaciness dengan nilai maksimum yang diperoleh dari keduanya (Mustofa, 2020). Adapun persamaan Silhouette Index dari sebuah cluster sebagai berikut (Prasetyo, 2014):

$$SI_{j} = \frac{1}{m_{j}} \sum_{i=1}^{m_{j}} SI_{i}^{j}$$
 (2.8)

Dimana:

 $SI_i^j = Silhouette Index$  data ke-i dalam cluster ke-j

 $m_i$  = jumlah data pada *cluster* ke-j

Nilai dari  $SI_i^j$  diperoleh dari persamaan berikut:

$$SI_{i}^{j} = \frac{b_{i}^{j} - a_{i}^{j}}{max\{b_{i}^{j}, a_{i}^{j}\}}$$
 (2.9)

Dimana:

 $b_i^j$  = rata-rata jarak data ke-i terhadap data lain di luar *cluster* j

 $a_i^j = \operatorname{rata-rata}$  jarak data ke-i terhadap data lain dalam  $\operatorname{cluster}$  j

Nilai dari  $b_i^j$  dan  $a_i^j$  diperoleh dari persamaan berikut:

$$a_i^j = \frac{1}{m_j^{-1}} \sum_{\substack{r=1\\r \neq i}}^{m_j} d(x_i^j, x_r^j)$$
 (2.10)

dan

$$b_{i}^{j} = \min_{\substack{n=1,\dots,k\\ n \neq j}} \left\{ \frac{1}{m_{n}} \sum_{\substack{r=1\\r \neq i}}^{m_{n}} d(x_{i}^{j}, x_{r}^{n}) \right\}$$
(2.11)

Dimana:

 $i = 1, 2, ..., m_j$ 

 $d(x_i^j,x_r^j) \ = \ \mathrm{jarak}$ data ke-i dengan data ke-r dalam satu  $\mathit{cluster}$  j

 $m_j$  = jumlah data dalam *cluster* ke-j

Nilai dari Silhouette Index secara global diperoleh dari perhitungan rata-rata

nilai Silhouette Index pada semua cluster dengan persamaan sebagai berikut:

$$SI = \frac{1}{k} \sum_{j=1}^{k} SI_j$$
 (2.12)

Dimana:

k = jumlah cluster

 $SI_j = Silhouette Index pada cluster ke-j$ 

Nilai yang didapat dari *Silhouette Index* berada pada rentang -1 sampai +1. Nilai *Silhouette Index* yang mendekati +1 menunjukkan data tidak sesuai dengan *cluster* yang ditempati. Data akan berada pada *cluster* yang tepat jika nilai dari *Silhouette Index* mendekati -1, sedangkan nilai *Silhouette Index* yang mendekati 0 menujukkan data berada di antara dua *cluster*. Menurut Kaufman dan Rousseeuw (2005) nilai ukur dari *Silhouette Index* sebagai berikut:

Tabel 2.1 nilai ukur Silhouette Index

Nilai Silhouette Index	Keterangan
0,71-1,00	Klaster yang kuat
0,51-0,70	Klaster telah layak atau sesuai
0,26-0,50	Klaster yang lemah
$\leq 0,25$	Tidak dapat dikatakan sebagai klaster

#### **BAB III**

#### METODE PENELITIAN

#### 3.1. Jenis Penelitian

Ditinjau dari segi paradigma penelitian, jenis penelitian yang dilakukan merupakan penelitian kuantitatif karena didasarkan pada penelitian sebelumnya. Berdasarkan sumber data, jenis penelitian ini termasuk dalam jenis penelitian sekunder, karena data diperoleh dari pihak ke tiga yaitu Badan Pusat Statistik Surabaya.

#### 3.2. Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang didapatkan dari hasil Survey Ekonomi Nasionan (SUSENAS) yang dilakukan oleh Badan Pusat Statistik Surabaya di tahun 2018. Data yang digunakan merupakan sampel dari rumah tangga yang berada di setiap kecematan dengan jumlah 1070 rumah tangga. Adapun data yang digunakan berupa luas bangunan per orang  $(m^2)$ , jenis lantai, jenis dinding, fasilitas buang air besar, melewatkan makan, makan lebih sedikit, kehabisan makanan, memiliki BPJS PBI, memiliki BPJS Non PBI, memiliki jamkesmas, memiliki asuransi kesehatan swasta, memiliki asuransi kesehatan perusahaan, tidak punya asuransi kesehatan, pendidikan terakhir kepala rumah tangga, lapangan pekerjaan kepala rumah tangga, memiliki kulkas, memiliki AC, memiliki mobil.

### 3.3. Variabel Penelitian

Variabel yang akan digunakan dalam penelitian ini akan ditunjukkan pada tabel 3.1.

**Tabel 3.1 Variabel Penelitian** 

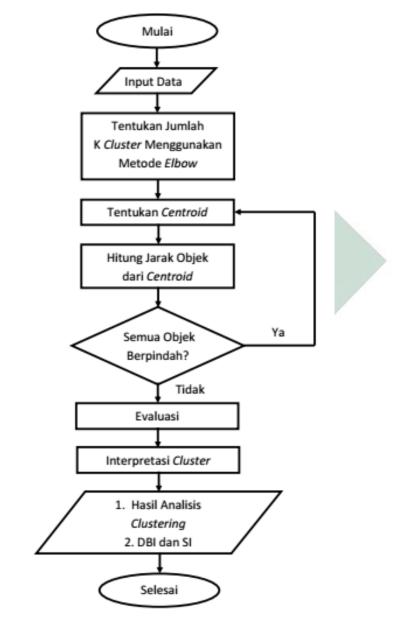
Kategori	Variabel	Keterangan	Label
Kondisi	x1	Luas Bangunan per orang	-
Bangunan	x2	Jenis Lantai	1 : Marmer/Granit
			2 : Keramik
			3 : Parket/Vinil/Karpet
		/ A	4 : Ubin/Tegel/Teraso
			5 : Kayu/Papan
			6 : Semen/Bata Merah
			7 : Bambu
			8 : Tanah
			9 : Lainnya
	х3	Jenis Dinding	1 : Tembok
			2 : Plasteran Anyaman Bambu/Kawat
			3 : Kayu/Papan
			4 : Anyaman Bambu
			5 : Batang Kayu
			6 : Bambu
			7 : Lainnya
	x4	Fasilitas Buang Air Besar	0 : Tidak Punya
			1 : Tangki
			2: IPAL

			3 : Kolam/Sawah/Sungai/Danau/Laut
			4 : Lubang Tanah
			5 : Pantai/Tanah Lapang/Kebun
			6 : Lainnya
Pangan	x5	Melewatkan Makan	1 : Iya
			5 : Tidak
			8 : Tidak Tahu
			9 : Menolak Menjawab
	x6	Makan Lebih Sedikit	1 : Iya
4		41	5 : Tidak
- 41		/ 2 / 3	8 : Tidak Tahu
			9 : Menolak Menjawab
	x7	Kehabisan Makanan	1 : Iya
			5 : Tidak
			8 : Tidak Tahu
			9 : Menolak Menjawab
Kesehatan	x8	Memiliki BPJS PBI	1 : Iya
			5 : Tidak
	x9	Memiliki BPJS Non PBI	1 : Iya
			5 : Tidak
	x10	Memiliki JAMKESMAS	1 : Iya
			5 : Tidak
	x11	Memiliki Asuransi	1 : Iya
		Kesehatan Swasta	5 : Tidak
	x12	Memiliki Asuransi	1 : Iya

		Kesehatan Perusahaan	5 : Tidak
	x13	Tidak Punya Asuransi	1 : Iya
		Kesehatan	5 : Tidak
Karakteristik	x14	Lapangan Pekerjaan	0 : Tidak memiliki lapangan pekerjaan
Kepala			1 : Pertanian Tanaman pada Palawija
Rumah Tangga			2 : Hortikultura
			3 : Perkebunan
			4 : Perikanan
		AL A	26 : Aktivitas Internasional
	x15	Pendidikan Terakhir	1 : Tidak Punya Ijazah SD
		Kepala Rumah Tangga	2 : Paket A
			3 : SDLB
			4 : SD
			5 : MI
			6 : SDLB
			7 : SMPLB
			21 : S3
Kepemilikan	x16	Memiliki Kulkas	1 : Iya
Aset			5 : Tidak
	x17	Memiliki Ac	1 : Iya
			5 : Tidak
	x18	Memiliki Mobil	1 : Iya
			5 : Tidak

## 3.4. Tahap Penelitian

Tahap penelitian yang akan dilakukan ditunjukkan dalam bentuk *flawchart* pada Gambar 3.1.



Gambar 3.1 Flawchart K-Means

Proses Penelitian yang akan dilakukan berdasarkan Gambar 3.1, sebagai berikut:

- 1. Memasukkan data pada setiap kategori yang telah ditentukan.
- 2. Menentukan jumlah K *cluster* pada setiap kategori yang akan digunakan, menggunakan metode *elbow* dengan Persamaan (2.3).
- 3. Menentukan *centroid* pada setiap kategori secara acak, untuk dilakukan proses pengklasteran menggunakan algoritma K-Means dengan Persamaan (2.1) sampai optimum.
- 4. Mengevaluasi hasil *cluster* pada setiap kategori yang didapat, menggunakan *silhouette index* dan *davies-bouldin index* dengan Persamaan (2.12) dan (2.4).
- 5. Interpretasi *cluster* pada setiap kategori

#### **BAB IV**

#### HASIL DAN PEMBAHASAN

### 4.1. Deskripsi Data

Penelitian ini menggunakan data yang diperoleh langsung dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) yang dilakukan oleh Badan Puasat Statistik Surabaya pada tahun 2018. Data berisi 18 variabel yang dikelompokkan berdasarkan 5 keategori yang terdiri dari kondisi bangunan, pangan, kesehatan, karakteristik kepala rumah tangga dan kepemilikan aset. Adapun kategori kondisi bangunan meliputi luas bangunan per orang, jenis lantai, jenis dinding, fasilitas buang air besar. Kategori pangan meliputi melewatkan makan, makan lebih sedikit dan kehabisan makanan. Kategori kesehatan meliputi memiliki BPJS PBI, memiliki BPJS Non PBI, memiliki Jamkesmas, memiliki asuransi kesehatan swasta, memiliki asuransi kesehatan perusahaan dan tidak punya asuransi kesehatan. Kategori karakteristik kepala rumah tangga meliputi pendidikan terakhir kepala rumah tangga dan lapangan pekerjaan kepala rumah tangga. Kategori kepemilikan aset meliputi memiliki kulkas, memiliki AC dan memiliki mobil. Jumlah data yang digunakan sebanyak 1070 rumah tangga yang merupakan sampel yang diambil dari setiap kecamatan.

#### 4.2. Menentukan Jumlah K Cluster Menggunakan Metode Elbow

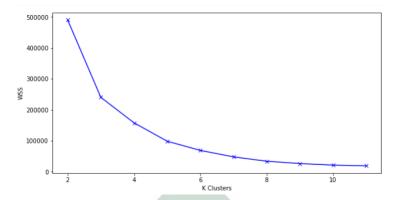
Metode *elbow* digunakan untuk menentukan jumlah K *cluster* optimum. Pada dasarnya, metode *elbow* merupakan metode WSS (*Within Cluster Sum of Square*) atau biasa disebut dengan SSE (*Sum Square Error*) yang ditunjukkan pada Persamaan 2.3. Penentuan K *cluster* optimum, dapat dilihat pada grafik hasil perhitungan

WSS yang berbentuk siku atau biasa disebut dengan *elbow point*. Biasanya, *elbow point* terletak setelah penurunan nilai WSS yang tinggi kemudian disertai dengan penurunan WSS yang mulai stabil setelahnya. Berikut merupakan hasil dari perhitungan *elbow* pada setiap kategori.

Pada kategori kondisi bangunan, nilai WSS dapat diamati pada Tabel 4.1 yang menunjukkan terjadinya penurunan yang tinggi sampai pada K=5 kemudian disertai penurunan mulai stabil setelahnya. Pada Grafik 4.1 dapat dilihat *elbow point* terletak pada K=5. Berdasarkan pengamatan tersebut dapat diambil kesimpulan, banyak K *cluster* optimum pada kategori kondisi bangunan yaitu K=5.

Tabel 4.1 WSS Pada Kategori Kondisi Bangunan

Clusters	wss
2	490 <mark>61</mark> 1.03 <mark>502</mark> 5
3	<mark>2</mark> 40287.912 <mark>41</mark> 9
4	157948.841911
5	98406.753773
6	68738.573172
7	47741.163490
8	33725.057571
9	26236.842006
10	21104.969238
11	18816.512333

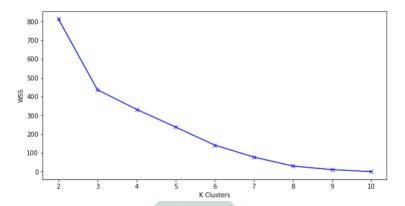


Gambar 4.1 Grafik WSS Pada Kategori Kondisi Bangunan

Pada kategori pangan, nilai WSS dapat diamati pada Tabel 4.2. Penurunan mulai stabil setelah berada pada K=3. Pada Grafik 4.2 juga menunjukkan *elbow point* terletak pada K=3. Dari pengamatan tersebut dapat diambil kesimpulan, banyak K *cluster* optimum pada kategori pangan yaitu K=3.

Tab<mark>el 4.2 WSS Pa</mark>da Kategor<mark>i P</mark>angan

Clusters	WSS
2	813.4197
3	435.6462
4	331.9383
5	237.3605
6	141.3584
7	77.55218
8	29.69803
9	10.66667
10	0.00

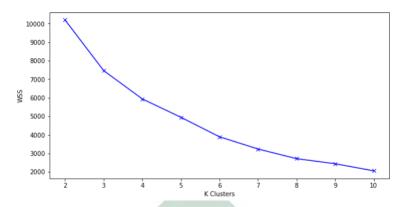


Gambar 4.2 Grafik WSS Pada Kategori Pangan

Pada kategori kesehatan, jika diamati melalui Grafik 4.3 *elbow point* terlihat berada pada titik K = 6. Sedangkan pada Tabel 4.3 penurunan nilai WSS terlihat masih tetap tinggi hingga nilai K = 6. Dari pengamatan tersebut dapat diambil kesimpulan, banyak K *cluster* optimum pada kategori kesehatan yaitu K = 6.

Tabel 4.3 WSS Pada Kategori Kesehatan

Clusters	wss
2	10196.549371
3	7464.946467
4	5940.790561
5	4949.849744
6	3898.036737
7	3237.882788
8	2716.129655
9	2442.997136
10	2059.168688

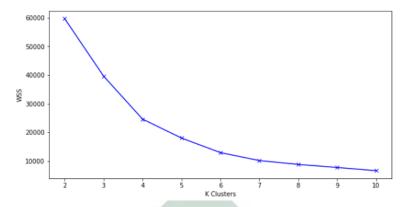


Gambar 4.3 Grafik WSS Pada Kategori Kesehatan

Pada kategori karakteristik kepala rumah tangga, nilai WSS dapat diamati pada Tabel 4.4 yang menunjukkan terjadinya penurunan yang tinggi sampai pada K = 6 kemudian disertai penurunan mulai stabil setelahnya. Pada Grafik 4.4 dapat dilihat *elbow point* terletak pada K = 6. Dari pengamatan tersebut dapat diambil kesimpulan, banyak K *cluster* optimum pada kategori kondisi bangunan yaitu K = 6.

Tabel 4.4 WSS Pada Kategori Karakteritik Kepala Rumah Tangga

Clusters	WSS	
2	59681.695677	
3	39617.153999	
4	24651.845858	
5	18042.307613	
6	12972.811019	
7	10192.718947	
8	8851.194078	
9	7783.245396	
10	6660.392364	

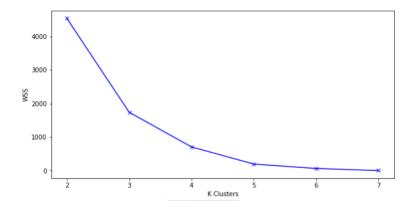


Gambar 4.4 Grafik WSS Pada Kategori Karakteritik Kepala Rumah Tangga

Pada kategori kepemilikan aset, penurunan nilai WSS pada Tabel 4.5 terlihat sangat tinggi sampai pada K = 4. Pada Grafik 4.5 dapat dilihat *elbow point* terletak pada K = 4. Dari pengamatan tersebut dapat diambil kesimpulan, banyak K *cluster* optimum pada kategori kondisi bangunan yaitu K = 4.

Tabel 4.5 WSS Pada Kategori Kepemilikan Aset

Clusters	wss
2	4541.988
3	1739.196
4	705.1005
5	194.3935
6	62.976
7	0.00



Gambar 4.5 Grafik WSS Pada Kategori Kepemilikan Aset

### 4.3. Perhitungan Analisis Cluster Menggunakan Metode K-Means

Setelah dilakukan pemilihan K optimum menggunakan metode *elbow* pada setiap kategori. Data kemudian dikelompokkan berdasarkan kemiripan karakteristik menggunakan algoritma K-Means *cluster*. Berikut merupakan contoh perhitungan manual K-Means *cluster* pada kategori kepemilikan aset. Adapun jumlah data yang digunakan sebanyak 10 data yang dipilih secara acak yang ditunjukkan pada Tabel 4.6.

Sebelum melakukan pengklasteran, terlebih dahulu melelakukan pemilihan *centroid* secara acak sebanyak K. Berdasarkan metode *elbow*, banyak K *cluster* yang digunakan yaitu 4. Adapun *centroid* yang digunakan terdapat pada Tabel 4.7 yang merupakan data ke-2, 3, 6 dan 7.

Tabel 4.6 Data Kepemilikan Aset

Data Ke-i	x16	x17	x18
1	1	5	5
2	1	5	5
3	5	5	5
4	1	1	5
5	5	5	5
6	5	5	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	5

Tabel 4.7 Centroid Pada Iterasi ke-1

Centroid	x16	x17	x18
c1	1	5	5
c2	5	5	5
c3	5	5	1
c4	1	1	1

Setelah menentukan *centroid* secara acak, kemudian menghitung jarak setiap data terhadap *centroid* menggunakan *Euclidean* pada Persamaan 2.1. Berikut merupakan contoh perhitungan jarak data ke-1 terhadap setiap *centroid*:

$$d(1,1) = \sqrt{\sum_{i=1}^{r} (x_i, y_i)^2}$$

$$= \sqrt{(1-1)^2 + (5-5)^2 + (5-5)^2}$$

$$= \sqrt{0}$$

$$= 0$$

$$d(1,2) = \sqrt{\sum_{i=1}^{r} (x_i, y_i)^2}$$

$$= \sqrt{(1-5)^2 + (5-5)^2 + (5-5)^2}$$

$$= \sqrt{16}$$

$$= 4$$

$$d(1,3) = \sqrt{\sum_{i=1}^{r} (x_i, y_i)^2}$$

$$= \sqrt{(1-5)^2 + (5-5)^2 + (5-1)^2}$$

$$= \sqrt{32}$$

$$= 5.656854249$$

$$d(1,4) = \sqrt{\sum_{i=1}^{r} (x_i, y_i)^2}$$

$$= \sqrt{(1-1)^2 + (5-1)^2 + (5-1)^2}$$

$$= \sqrt{32}$$

$$= 5.656854249$$

Tabel 4.8 Hasil Perhitungan Jarak Pada Iterasi ke-1

Data	Jarak ke centroid				Cluster
ke-i	c1	c2	c3	c4	yang diikuti
1	0	4	5.656854249	5.656854249	1
2	0	4	5.656854249	5.656854249	1
3	4	0	4	6.92820323	2
4	4	5.656854249	6.92820323	4	1
5	4	0	4	6.92820323	2
6	5.656854249	4	0	5.656854249	3
7	5.656854249	6.92820323	5.656854249	0	4
8	5.656854249	6.9 <mark>2</mark> 820323	5.656854249	0	4
9	5.656854249	6.92820 <mark>32</mark> 3	5.6 <mark>568</mark> 54249	0	4
10	4	<mark>5.</mark> 656 <mark>85424</mark> 9	6.9 <mark>282</mark> 0323	4	1

Setelah menghitung jarak setiap data, kemudian menentukan rata-rata setiap data yang bergabung pada *cluster* yang diikuti. Perhitungan rata-rata ini bertujuan untuk menentukan, apakah *centoid* pada iterasi ke-1 mengalami perubahan atau tidak. Berikut merupakan perolehan rata-rata pada setiap *cluster*:

Tabel 4.9 Rata-Rata Cluster ke-1 Pada Iterasi ke-1

Data k-i	x16	x17	x18
1	1	5	5
2	1	5	5
4	1	1	5
10	1	1	5
Rata-rata	1	3	5

Tabel 4.10 Rata-Rata Cluster ke-2 Pada Iterasi ke-1

Data ke-i	x16	x17	x18
3	5	5	5
5	5	5	5
Rata-rata	5	5	5

Tabel 4.11 Rata-Rata Cluster ke-3 Pada Iterasi ke-1

Data kw-i	x16	x17	x18
6	5	5	1
Rata-rata	5	5	1

Tabel 4.12 Rata-Rata *Cluster* ke-4 Pada Iterasi ke-1

<mark>Dat</mark> a k <mark>e-i</mark>	x16	x17	x18
7	1	1	1
8	1	1	1
9	1	1	1
Rata-rata	_1	1	1

Dilihat dari nilai rata-rata pada setiap *cluster* yang diperoleh, terdapat perubahan *centroid* pada *cluster* ke-1. Untuk memastikan apakah terdapat objek yang berpindah atau tidak pada setiap *cluster*, maka diperlukan pengklasteran pada iterasi ke-2 menggunakan *centroid* baru yang ditunjukkan pada Tabel 4.13.

Tabel 4.13 Centroid Pada Iterasi ke-1

Centroid	x16	x17	x18
c1	1	3	5
c2	5	5	5
c3	5	5	1
c4	1	1	1

dengan cara sama, pengklasteran yang dilakukan pada iterasi ke-2 menghasilkan nilai sebagaimana pada Tabel 4.14.

Tabel 4.14 Hasil Perhitungan Jarak Pada Iterasi ke-2

Data		Jara <mark>k k</mark> e <mark>centroid</mark>			Cluster
ke-i	c1	c2	c3	c4	yang diikuti
1	2	4	5.65 <mark>68</mark> 54249	5.656854249	1
2	2	4	5.656854249	5.656854249	1
3	4.472135955	0	4	6.92820323	2
4	2	5.656854249	6.92820323	4	1
5	4.472135955	0	4	6.92820323	2
6	6	4	0	5.656854249	3
7	4.472135955	6.92820323	5.656854249	0	4
8	4.472135955	6.92820323	5.656854249	0	4
9	4.472135955	6.92820323	5.656854249	0	4
10	2	5.656854249	6.92820323	4	1

Pada iterasi ke-2, objek pada setiap *cluster* tidak mengalami perubahan. Oleh sebab itu, proses iterasi berhenti dengan hasil akhir *centroid* yang terdapat pada Tabel 4.13. Pada *centroid* dapat diketahui rumah tangga pada kategori sangat

miskin terdapat pada *cluster* ke-2 dengan rata-rata tidak memiliki aset berupa kul-kas, AC dan mobil. Rumah tangga dengan kategori miskin terdapat pada *cluster* ke-1 dengan rata-rata memiliki aset berupa kulkas, sebagian memiliki AC dan tidak memiliki mobil. Rumah tangga dengan kategori kaya terdapat pada *cluster* ke-3 dengan rata-rata tidak memiliki aset berupa kulkas dan AC, namun memiliki mobil. Rumah tangga dengan kategori sangat kaya terdapat pada *cluster* ke-3 dengan rata-rata memiliki aset berupa kulkas, AC dan mobil.

#### 4.4. Perhitungan Evaluasi Cluster

Evaluasi *cluster* dilakukan untuk menguji seberapa kuat *cluster* yang telah terbentuk pada setiap kategori yang telah ditentukan. Adapun algoritma yang digunakan yaitu *Silhouette Index* (SI) dan *Davies Bouldin Index* (DBI). Berikut merupakan contoh perhitungan manual dari evaluasi *cluster* menggunakan hasil *cluster* pada Tabel 4.14 dengan *centroid* optimum yang telah didapat pada Tabel 4.13.

#### 1. Davies Bouldin Index (DBI)

Evaluasi *clustering* menggunakan DBI dihitung menggunakan Persamaan 2.4. Adapun hal pertama yang perlu dilakukan untuk mencari nilai DBI, yaitu dengan mencari jarak *Euclidean* antara data dengan *centroid* yang diikuti. Berikut merupakan contoh perhitungan jarak dalam cluster ke-1:

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

$$= 2$$

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

$$= 2$$

$$d(4,1) = \sqrt{(1-1)^2 + (1-3)^2 + (5-5)^2}$$

$$= 2$$

$$d(10,1) = \sqrt{(1-1)^2 + (1-3)^2 + (5-5)^2}$$
$$= 2$$

Setelah menghitung jarak dari setiap data terhadap *centroid* yang diikutinya, kemudian menentukan nilai SSW pada setiap *cluster* menggunakan Persamaan 2.6. Berikut merupakan contoh perhitungan SSW pada *cluster* ke-1:

$$SSW_1 = \frac{1}{4}(2+2+2+2) = 2$$

Hasil perhitungan SSW pada setiap cluster disajikan Pada Tabel 4.15.

Tabel 4.15 Hasil SSW Pada Setiap Cluster

Data	- 4	variabe	1	Cluster	Centroid		Jarak	SSW	
ke-i	x16	x17	x18	yang <mark>d</mark> iikuti	x16	x17	x18	ke Centroid	
1	1	5	5	1	1	3	5	2	2
2	1	5	5	1				2	
4	1	1	5	1		1		2	
10	1	1	5	1				2	
3	5	5	5	2	5	5	5	0	0
5	5	5	5	2				0	
6	5	5	1	3	5	5	1	0	0
7	1	1	1	4	1	1	1	0	0
8	1	1	1	4				0	
9	1	1	1	4				0	

Setelah mendapatkan nilai SSW dari setiap *cluster*, kemudian menghitung jarak antar *centroid cluster* atau bisa disebut SSB dengan menggunakan Persamaan 2.7 sebagai berikut:

$$SSB_{1,2} = d(c_1, c_2)$$

$$= \sqrt{(1-5)^2 + (3-5)^2 + (5-5)^2}$$

$$= \sqrt{20}$$

$$= 4.472135955$$

$$SSB_{1,3} = d(c_1, c_3)$$

$$= \sqrt{(1-5)^2 + (3-5)^2 + (5-1)^2}$$

$$= \sqrt{36}$$

$$= 6$$

$$SSB_{1,4} = d(c_1, c_4)$$

$$= \sqrt{(1-1)^2 + (3-1)^2 + (5-1)^1}$$

$$= \sqrt{20}$$

$$= 4.472135955$$

Hasil perhitungan SSB disajikan pada Tabel 4.16.

Tabel 4.16 Hasil SSB Antar Centroid

Centroid ke-i	1	2	3	4
1	0	4.472135955	6	4.472135955
2	4.472135955	0	4	6.92820323
3	6	4	0	5.656854249
4	4.472135955	6.92820323	5.656854249	0

Setelah mendapatkan nilai dari SSB, kemudian mencari nilai dari rasio antar *cluster* menggunakan Persamaan 2.5 sebagai berikut:

$$R_{1,2} = \frac{SSW_1 + SSW_2}{SSB_{1,2}}$$

$$= \frac{2+0}{4.472135955}$$

$$= 0.447213595$$

$$R_{1,3} = \frac{SSW_1 + SSW_3}{SSB_{1,3}}$$

$$= \frac{2+0}{6}$$

$$= 0$$

$$R_{1,4} = \frac{SSW_1 + SSW_4}{SSB_{1,4}}$$

$$= \frac{2+0}{4.472135955}$$

$$= 0.447213595$$

pada perhitungan rasio diatas diketahui  $R_{max}$  pada cluster ke-1 yaitu 0.447213595. Dengan cara yang sama pada *cluster* lain, perhitungan rasio dilakukan sampai didapatkan  $R_{max}$ . Jika nilai  $R_{max}$  dimasukkan dalam rumus DBI yang tertera pada Persamaa 2.4, maka akan menghasilkan nilai DBI sebagai berikut:

$$DBI = \frac{1}{4}(0.447213595 + 0.447213595 + 0 + 0.447213595)$$
$$= 0.335410197$$

Hasil perhitungan nilai DBI yang didapat pada contoh tersebut adalah 0.335410197. Nilai DBI yang mendekati 0 menunjukkan hasil *cluster* yang baik.

#### 2. Silhouette Index (SI)

Nilai SI pada keseluruhan hasil cluster (SI global) dihitung menggunakan Persamaan 2.12. Adapun hal pertama yang dilakukan yaitu mencari nilai SI pada setiap cluster dengan mencari nilai rata-rata SI pada setiap data. Untuk mencari nilai SI pada data, dibutuhkan nilai  $a_i^j$  dan  $b_i^j$  yang didapat dari Persamaan 2.10 dan 2.11. Berikut merupakan contoh perhitungan  $a_i^j$  pada data yang berada di cluster

$$\begin{array}{ll} a_1^1 &=& \frac{1}{4-1}(\sqrt{(1-1)^2+(5-5)^2+(5-5)^2}+\sqrt{(1-1)^2+(5-1)^2+(5-1)^2}+\\ && \sqrt{(1-1)^2+(5-1)^2+(5-5)^2})\\ &=& \frac{1}{3}(0+4+4)\\ &=& 2.666666667\\ a_2^1 &=& \frac{1}{4-1}(\sqrt{(1-1)^2+(5-5)^2+(5-5)^2}+\sqrt{(1-1)^2+(5-1)^2+(5-1)^2}+\\ && \sqrt{(1-1)^2+(5-1)^2+(5-5)^2})\\ &=& \frac{1}{3}(0+4+4)\\ &=& 2.666666667\\ a_4^1 &=& \frac{1}{4-1}(\sqrt{(1-1)^2+(1-5)^2+(5-5)^2}+\sqrt{(1-1)^2+(1-5)^2+(5-5)^2}+\\ && \sqrt{(1-1)^2+(1-5)^2+(5-5)^2})\\ &=& \frac{1}{3}(4+4+0)\\ &=& 2.666666667\\ a_{10}^1 &=& \frac{1}{4-1}(\sqrt{(1-1)^2+(1-5)^2+(5-5)^2}+\sqrt{(1-1)^2+(1-5)^2+(5-5)^2}+\\ && \sqrt{(1-1)^2+(1-5)^2+(5-5)^2})\\ &=& \frac{1}{3}(4+4+0)\\ &=& 2.6666666667\\ \end{array}$$

untuk perhitungan  $b_i^j$  pada data yang berada pada *cluster* ke-1 sebagai beri-

kut:

$$b_1^1 = min\left\{\frac{1}{2}(4+4), 5.656854249, \frac{1}{3}(5.656854249 + 5.656854249 + 5.656854249)\right\}$$
$$= min(4, 5.656854249, 5.656854249)$$
$$= 4$$

$$\begin{array}{ll} b_1^2 &=& \min \left\{ \frac{1}{2} (4+4), 5.656854249, \frac{1}{3} (5.656854249 + 5.656854249 + 5.656854249) \right. \\ &=& \min (4, 5.656854249, 5.656854249) \\ &=& 4 \\ b_1^4 &=& \min \left\{ \frac{1}{2} (5.656854249 + 5.656854249), 6.92820323, \frac{1}{3} (4+4+4) \right\} \\ &=& \min (5.656854249, 6.92820323, 4) \\ &=& 4 \\ b_1^{10} &=& \min \left\{ \frac{1}{2} (5.656854249 + 5.656854249), 6.92820323, \frac{1}{3} (4+4+4) \right\} \\ &=& \min (5.656854249, 6.92820323, 4) \\ &=& 4 \end{array}$$

Untuk perhitungan SI pada data yang berada pada *cluster* ke-1 sebagai berikut:

$$SI_1^1 = \frac{4-2.666666667}{max\{4,2.666666667\}} = 0.1333333333$$
  
 $SI_2^1 = \frac{4-2.666666667}{max\{4,2.666666667\}} = 0.1333333333$   
 $SI_4^1 = \frac{4-2.666666667}{max\{4,2.666666667\}} = 0.1333333333$   
 $SI_{10}^1 = \frac{4-2.666666667}{max\{4,2.666666667\}} = 0.1333333333$ 

Untuk perhitungan SI pada cluster ke-1 sebagai berikut:

$$SI_1 = \frac{1}{m_1} (SI_1^1 + SI_2^1 + SI_4^1 + SI_{10}^1)$$
  
=  $\frac{1}{4} (0.1333333333 + 0.133333333 + 0.133333333 + 0.133333333))$   
=  $0.1333333333$ 

dengan cara yang sama, SI pada setiap *cluster* dihitung. Setelah mendapatkan nilai SI pada setiap *cluster*, kemudian menghitung SI global menggunaakan Persamaan 2.12, sebagai berikut:

Hasil nilai SI yang didapat pada contoh tersebut adalah 0.783333333. Hal tersebut dapat diartikan bahwa, setiap data telah masuk pada *cluster* yang tepat.

Adapun hasil evaluasi *cluster* berdasarkan data asli pada setiap kategori pengelompokan, menggunakan metode DBI dan SI dapat dilihat pada Tabel 4,17.

Kategori **DBI** SI Keterangan 1 0.467166217 0.611188336 Klaster layak atau sesuai 2 0.960980645 0.33384507 Klaster kuat 3 0.653452852 0.656813758 Klaster layak atau sesuai 4 0.607747038 0.539465257 Klaster layak atau sesuai 5 0.284760075 0.919161316 Klaster kuat

Tabel 4.17 Hasil Evaluasi Cluster Pada Setiap Kategori

Pada Tabel 4.17, dapat diketahui bahwa hasil evaluasi pengelompokan pada setiap kategori memiliki nilai *silhouette index* lebih dari 0.50 serta memiliki nilai *davies bouldin index* mendekati 0. Hal tersebut menenjukkan, pengelompokan pada setiap kategori dapat digunakan.

## 4.5. Interpretasi *Cluster*

Interpretasi *cluster* pada kategori bangunan, pangan, kesehatan, karakteristik kepala rumah tangga dan kepemilikan aset dilakukan untuk mengetahui gambaran mengenai karakteristik pada rumah tangga dan untuk mengetahui jumlah rumah tangga yang masuk pada kategori miskin. Adapun penggambaran karakteristik

yang dilakukan merupakan peninjauan berdasarkan pada pada hasil perhitungan K-Means *cluster* yang dioptimasi menggunakan metode *elbow* yang telah dilakukan sebelumnya.

### 4.5.1. Kategori Kondisi Bangunan

Perhitungan K-Means *cluster* pada kategori kondisi bangunan dengan penentuan K *cluster* menggunakan metode *elbow* yaitu K = 5. Adapun nilai variabel yang diperoleh pada setiap *cluster* dipengaruhi oleh label yang telah disebutkan pada Tabel 3.1. Nilai pada variabel x1 menunjukkan, semakin tinggi nilai maka bangunan rumah tangga pada suatu *cluster* tersebut semakin luas. Nilai pada variabel x2, x3 dan x4 menunjukkan semakin tinggi nilai variabel yang didapat pada suatu *cluster*, maka bahan bangunanan yang digunakan dapat dikatakan kurang layak. Untuk mengetahui lebih jelas karakteristik kondisi bangunan pada setiap *cluster*, dapat dilihat pada Tabel 4.18 mengenai *centroid* dari hasil pengklasteran yang telah didapat.

Tabel 4.18 Hasil Centroid Cluster Pada Kategori Kondisi Bangunan

Centroid Cluster ke-i	Jumlah Rumah Tangga	x1	x2	х3	x4
1	97	76.06	2.34	1.08	1
2	656	9.74	2.621	1.16	1.21
3	27	159.60	1.96	1	1.04
4	3	313	2	1	1
5	287	33.47	2.25	1.03	1.02

Pada Tabel 4.18, karakteristik rumah tangga yang masuk dalam *cluster* 1 rata-rata memiliki luas bangunan per orang 76.06  $m^2$ , jenis lantai yang digunak-

an mayoritas keramik, jenis dinding mayoritas berupa tembok dan setiap rumah tangga yang tergolong didalamnya menggunakan fasilitas buang air besar berupa tangki. Rata-rata karakterisik rumah tangga yang masuk dalam cluster 2 memiliki luas bangunan per orang 9.74  $m^2$ , jenis lantai yang digunakan berupa parket/vini-l/karpet, dinding pada bangunan berupa tembok dan menggunakan fasilitisas buang air besar menggunakan tangki. Pada cluster 3 terdapat rumah tangga yang memiliki karakteristik dengan rata-rata luas bangunan per orang 159.6  $m^2$ , jenis lantai yang digunakan berupa keramik dan menggunakan fasilitas buang air besar mengunakan tangki. Adapun jenis tembok yang digunakan, keseluruhan rumah tangga yang masuk di dalamnya yaitu berupa tembok. Rumah tangga yang masuk pada cluster 4 rata-rata memiliki karakteristik luas bangunan per orang 313  $m^2$  dan keseluruhan rumah tangga menggunakan jenis lantai berupa keramik, dinding bangunan berupa tembok, serta fasilitas buang air besa rberupa tangki. Pada cluster 5 rata-rata rumah tangga memiliki luas bangunan 33.5  $m^2$ , menggunakan jenis lantai berupa keramik, jenis dinding berupa tembok, serta fasilitas buang air besar menggunakan tangki.

Berdasarkan karakteristik rumah tangga yang telah diketahui pada setiap *cluster*, maka rumah tangga yang telah dikelompokkan tersebut dapat dikategorikan tingkat kemiskinan menurut karakteristik bangunan sebagai berikut:

Tabel 4.19 Hasil Cluster Pada Kategori Kondisi Bangunan

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Sangat Miskin	2	656
Miskin	5	287
Cukup	1	97
Kaya	3	27
Sangat Kaya	4	3

### 4.5.2. Kategori Pangan

Hasil dari K-Means *cluster* yang didapatkan pada kategori pangan, dengan pemilihan K *cluster* sesuai dengan metode *elbow* yang diterapkan yaitu K = 3 diperoleh nilai *centroid* pada setiap *cluster* sebagaimana yang tertera pada Tabel 4.20. Nilai *centroid* pada variabel x5, x6 dan x7 yang semakin rendah menunjukkan kondisi pangan pada rumah tangga yang masuk dalam ketegori tersebut masih bisa dikatakan mampu untuk mencukupi kebutuhan pangan.

Tabel 4.20 Hasil Centroid Pada Kategori Pangan

Centroid		Jumlah	x5	х6	x7
Clus	ster ke-i	Rumah Tangga			
	1	37	4.24	1	4.35
	2	995	4 <mark>.99</mark>	5	4.99
	3	38	1	1	1

Pada Tabel 4.20 menunjukkan karakteristik pada setiap *cluster*. Dimana *cluster* 1 menunjukkan rumah tangga yang masuk pada kelompok tersebut mayoritas tidak pernah melewatkan makan, pernah makan lebih sedikit dan mayoritas pernah kehabisan makanan. Rumah tangga pada *cluster* 2 memiliki karakteristik tidak pernah melewatkan makan, tidak pernah makan lebih sedikit dan tidak pernah kehabisan makan. Rumah tangga pada *cluster* 3 pernah melewatkan makan, pernah makan lebih sedikit dan pernah kehabisan makan. Berdasarkan karakteristik tersebut dapat diketahui tingkat kemiskinan berdasarkan kategori pangan sebagai berikut:

Tabel 4.21 Hasil Cluster Pada Kategori Pangan

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Miskin	3	38
Cukup	1	37
Kaya	2	995

### 4.5.3. Kategori Kesehatan

Pengelompokan rumah tangga miskin berdasarkan pada hasil K-Means *cluster*, dengan pemilihan K *cluster* menggunkana metode *elbow* didapat hasil *centroid* pada tabel 4.22. Pada kategori ini, kesehatan suatu rumah tangga dilihat melalui asuransi kesehatan yang dimiliki. Adapun nilai variabel 0 yang didapat pada suatau *cluster* menunjukkan, bahwasannya rumah tangga yang berada *cluster* tersebut tidak memiliki asuransi kesehatan tersebut. Adapun keterangan setiap variabel dapat dilihat pada Tabel 3.1.

Tabel 4.22 Hasil Centroid Pada Kategori Kesehatan

Centroid	Jumlah	x8	x9	x10	x11	x12	x13
Cluster ke-i	Rumah Tangga						
1	195	1.8	4.30	4.16	4.53	4.96	5
2	101	4.76	1	4.92	4.64	4.72	1
3	86	4.63	5	4.91	4.86	1	4.02
4	290	5	1	4.88	4.81	4.67	5
5	313	5	5	4.88	4.80	5	1
6	85	1	5	4.95	5	4.39	1

Pada Tabel 4.22 dapat diketahui karakteristik rumah tangga berdasarkan ka-

tegori kesehatan pada setiap *cluster*. Adapun rumah tangga yang tergolong pada *cluster* 1 memiliki karakteristik dengan mayoritas memiliki asuransi kesehatan berupa BPJS PBI. Rumah tangga pada *cluster* 2 memiliki karakteristik dengan mayoritas memiliki BPJS PBI dan BPJS Non PBI, akan tetapi masih terdapat anggota rumah tangga yang tidak memiliki asuransi kesehatan. Pada *Cluster* 3 mayoritas rumah tangga hanya memiliki asuransi kesehatan perusahaan. Rumah tangga pada *cluster* 4 rata-rata hanya memiliki BPJS Non PBI. Rata-rata karakteristik rumah tangga pada *cluster* 4 tidak memiliki asuransi kesehatan. Pada *cluster* 6 rata-rata rumah tangga memiliki BPJS PBI, akan tetapi masih terdapat anggota rumah tangga yang tidak memiliki asuransi kesehatan. Berdasarkan karakteristik tersebut dapat diketahui tingkat kemiskinan berdasarkan kategori kesehatan sebagai berikut:

Tabel 4.23 Hasil Cluster Pada Kategori Kesehatan

Kategori	Cluster ke-i	J <mark>um</mark> lah Rumah Tangga
Sangat Sangat Miskin	5	313
Sangat Miskin	6	85
Miskin	1	195
Kaya	2	101
Sangat Kaya	4	290
Sangat Sangat Kaya	3	86

## 4.5.4. Kategori Karakteristik Kepala Rumah Tangga

Perhitungan K-Means cluster pada Karakteristik kepala rumah tangga dengan penentuan K *cluster* menggunakan metode *elbow*, yaitu K = 6 didapatkan *centroid* pada Tabel 4.24 Semakin rendah nilai variabel pada x15, menunjukkan semakin rendah tingkat pendidikan kepala rumah tangga tersebut. Adapun tinggi atau

rendahnya nilai variabel pada x14 dipengaruhi oleh pelabelan yang telah ditetapkan pada Tabel 3.1.

Tabel 4.24 Hasil Centroid Pada Kategori Karakterisritik Kepala Rumah Tagga

Centroid	Jumlah	x14	x15
Cluster ke-i	Rumah Tangga		
1	295	11.55	14.34
2	101	0.47	12.10
3	116	23.44	4.30
4	324	11.61	4.35
5	115	22.2	15.42
6	119	0.24	2.07

Pada Tabel 4.24 dapat diketahui gambaran rumah tangga berdasarkan kategori karakteristik kepala rumah tangga pada setiap *cluster*. Adapun rumah tangga yang tergolong pada *cluster* 1 memiliki karakteristik dengan mayoritas perdagangan besar dan eceran, reparasi perawatan mobil dan sepeda motor. Rata-rata pendidikan terakhir kepala rumah tangga adalah SMK. Pada *cluster* 2 mayoritas kepala rumah tangga tidak memiliki lapangan pekerjaan. Mayoritas pendidikan terakhir kepala rumah tangga adalah SMA. Pada *cluster* 3 mayoritas kepala rumah tangga memiliki aktivititas jasa lainnya. Mayoritas pendidikan kepala rumah tangga adalah SD. Pada *cluster* 4 mayoritas kepala rumah tangga memiliki lapangan pekerjaan berupa perdagangan besar dan eceran, reparasi, perawatan mobil dan sepeda motor. Mayoritas pendidikan kepala rumah tangga adalah SD. Pada *cluster* 5 mayoritas kepala rumah tangga memiliki aktivititas jasa lainnya. Mayoritas pendidikan kepala rumah tangga adalah SMA. Pada *cluster* 6 mayoritas kepala rumah tangga tidak memiliki lapangan pekerjaan. Mayoritas kepala rumah tangga tidak memiliki ijazah SD.

Berdasarkan karakteristik tersebut dapat diketahui tingkat kemiskinan berdasarkan kategori karakteristik kepala rumah tangga sebagai berikut:

Tabel 4.25 Hasil Centroid Pada Kategori Karakteristik Kepala Rumah Tangga

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Sangat Sangat Miskin	6	119
Sangat Miskin	2	101
Miskin	3	116
Kaya	4	324
Sangat Kaya	1	295
Sangat Sangat Kaya	5	115

# 4.5.5. Kategori Kepemilikan Aset

Pengelompokan rumah tangga miskin berdasarkan pada hasil K-Means *cluster*, dengan pemilihan K *cluster* menggunkana metode *elbow* didapat hasil *centroid* pada Tabel 4.26. Pada kategori ini, aset kepemilikan suatu rumah tangga dilihat melalui kepemilikan kulkas, mobil dan AC. Adapun nilai variabel 1 yang didapat pada suatau *cluster* menunjukkan, bahwasannya rumah tangga yang berada pada *cluster* tersebut memiliki aset sesuai pada variabel yang tertera, sedangkan nilai 5 menunjukkan rumah tangga yang berada dalam *cluster* tersebut tidak memiliki aset yang telah ditentukan. Keterangan setiap variabel dapat dilihat pada Tabel 3.1.

Tabel 4.26 Hasil Centroid Pada Kategori Kepemilikan Aset

Centroid	Jumlah	x16	x17	x18
Cluster ke-i	Rumah Tangga			
1	198	1	1.80	1
2	519	1	5	5
3	259	5	4.86	4.94
4	94	1	1	5

Pada Tabel 4.26 dapat diketahui karakteristik rumah tangga berdasarkan kategori kepemilikan aset. Adapun karakteristik rumah tangga yang tergolong pada *cluster* 1 rata-rata memiliki kulkas, mayoritas memiliki AC dan memiliki mobil. Pada *cluster* 2 rata-rata rumah tangga memiliki kulkas, tidak memiliki AC dan tidak memiliki mobil. Rumah tangga pada *cluster* 3 tidak memiliki kulkas, mayoritas tidak memiliki AC dan mayoritas tidak memiliki mobil. Pada *cluster* 4 rata-rata rumah tangga memiliki kulkas memiliki AC dan tidak memiliki mobil. Berdasarkan karakteristik tersebut dapat diketahui tingkat kemiskinan berdasarkan kategori kepemilikan aset sebagai berikut:

Tabel 4.27 Hasil Centroid Pada Kategori Kepemilikan Aset

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Sangat Miskin	3	259
Miskin	2	519
Kaya	4	94
Sangat Kaya	1	198

#### 4.6. Integrasi Keislaman

Kemiskinan merupakan permasalahan serius yang harus segera ditangani, terlebih pada saat pandemi seperti saat ini. Sebagaimana khalifah Ali bin bu Thaib mengatakan "Seandainya kefakiran/kemiskinan itu berwujud manusia, niscaya akan kubunuh dia". Kalimat tersebut menunjukkan betapa seriusnya pemerintahan Islam pada saat itu dalam menangani permasalahan kemiskinan. Jika kemiskinan terus meningkat, kemungkinan buruk yang akan dialami yaitu terjadinya kelumpuhan perkonomian yang menyebabkan permasalahan lain seperti tindak kejahatan kriminal juga akan meningkat. Untuk menghindari kemungkinan buruk tersebut dapat terjadi, pemerintah telah berupaya mengatasi permasalah kemiskinan tersebut. Salah satu upaya pemerintah dalam memberikan bantuan kepada masyarakat miskin merupakan langkah baik untuk membantu memenuhi kebutuhan mereka yang membutuhkan. Sebagaimana dalam dalam Al-Quran dijelaskan mengenai kemiskinan dalam surat Al-Baqarah ayat 273:

Artinya: "(Apa yang kamu infakkan) adalah untuk orang-orang fakir yang terhalang (usahanya karena jihad) di jalan Allah, sehingga dia yang tidak dapat berusaha di bumi; (orang lain) yang tidak tahu, menyangka bahwa mereka adalah orang-orang kaya karena mereka menjaga diri (dari meminta-minta). Engkau (Muhammad) mengenal mereka dari ciri-cirinya, mereka tidak meminta secara paksa kepada orang lain. Apapun harta yang baik yang kamu infakkan, sungguh, Allah Maha Mengetahui."

Permasalahan lain yang masih bersangkutan dengan kemiskinan yaitu keti-

daktepatan dalam meberikan bantuan. Padahal, ketepatan bantuan yang diberikan pemerintah merupakan suatu amanat yang seharusnya disampaikan. Sebagaimana hadist Rusullah SAW:

Artinya: ... dari Anas, ia berkata Rasulullah tidak berkhutbah kecuali bersabda: tidak ada keimanan bagi orang yang tidak melaksanakan amanah, dan tidak ada agama bagi orang yang tidak menepati janji. (H.R alThabarani)

Hadist tersebut menjelaskan, betapa pentingnya suatu amanat yang diemban, sampai-sampai dikatakan orang yang tidak melaksakan amanat berarti telah tidak beragama. Berdasarkan kenyataan yang dialami mengenai ketidak tepatan bantuan, menunjukkan pemerintahan yang kurang baik dalam mengemban amanat. Padahal, dalam hal seperti ini peran pemerintah mengenai kebijakan yang dibuat sangatlah penting. Sebagaimana dalam kaidah fiqih:

Artinya: Kebijaksanaan Imam/Kepala Negara terhadap rakyat itu harus dihubungkan dengan kemaslahatan.

Berdasarkan pada kaidah fiqih tersebut, kebijakan pemerintah sangatlah penting. Pemerintah seharusnya tidak hanya sekedar memberi bantuan, akan tetapi memastikan bahwa bantuan tersebut tidak salah sasaran juga tidak kalah penting. Dengan adanya penelitian ini yang membahas pengelompokan rumah tangga miskin sesuai kategori, bantuan tersebut dapat disalurkan sesuai dengan kebutuhan rumah tangga serta dapat mempercepat program pengentasan kemiskinan yang dilakukan pemerintah.

# **BAB V**

### **PENUTUP**

#### 5.1. Simpulan

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

- 1. Berdasarkan hasil optimasi metode K-Means clustering menggunakan metode elbow, diperoleh rumah tangga miskin yang pada setiap kategori. Pada kategori kondisi bangunan, diperoleh cluster rumah tangga sangat miskin sebanyak 656 dan rumah tangga miskin sebanyak 287. Pada kategori pangan terdapat rumah tangga miskin sebanyak 38. Pada kategori kesehatan didapat rumah tangga sangat-sangat miskin sebanyak 313, rumah tangga sangat miskin sebanyak 85 dan rumah tangga miskin sebanyak 195. Pada kategori karakteristik kepala rumah tangga diperoleh rumah tangga sangat-sangat miskin sebanyak 119, rumah tangga sangat miskin sebanyak 101 dan rumah tangga miskin sebanyak 116. Pada kategori kepemilikan aset, diperoleh rumah tangga sangat miskin sebanyak 259 dan rumah tangga miskin sebanyak 519.
- 2. Hasil validasi pengelompokan rumah tangga miskin di Surabaya menggunakan analisis *silhouette index* dan *davies-bouldin index* pada setiap kategori, diperoleh nilai *davies-bouldin index* sebesar 0.47 dan *silhouette index* sebesar 0.61 pada kategori kondisi bangunan. Pada kategori pangan, diperoleh

nilai davies-bouldin index sebesar 0.33 dan silhouette index sebesar 0.96. Pada kategori kesehatan, diperoleh nilai davies-bouldin index sebesar 0.65 dan silhouette index sebesar 0.65. Pada kategori karakteristik rumah tangga, diperoleh nilai davies-bouldin index sebesar 0.61 dan silhouette index sebesar 0.53. Pada kategori kepemilikian aset, diperoleh nilai davies-bouldin index sebesar 0.29 dan silhouette index sebesar 0.92. Hasil validasi berdasarkan silhouette index yang lebih dari 0.5 dan davies-bouldin index yang mendekati 0 pada setiap kategori menunjukkan cluster yang terbetuk dalam penelitian ini layak digunakan.

#### 5.2. Saran

Saran untuk penelitian lebih lanjut, dapat dilakukan penambahan variabel seperti jumlah penghasilan dan jumlah pengeluaran rumah tangga dalam satu bulan untuk menghasilkan pengelompokan rumah tangga miskin yang lebih akurat. Untuk memperoleh hasil yang lebih optimal dapat menggunakan metode *clustering* lainnya seperti K-Means++ dan K-Medoid.

### **DAFTAR PUSTAKA**

- Badan Statistik Penduduk Pusat Persentase Mis-2020 9,78%,15 2020, kin Maret Naik Menjadi Juli https://WWW.bps.go.id/pressrelease/2020/07/15/1744/persentase-pendudukmiskin-maret-2020-naik-menjadi-9-78-persen.html
- CNBC Indonesia *Sri Mulyani Bicara Dampak PSBB Luar Biasa Serius*,11 September 2020, https/WWW.cnbcindonesia.com/news/20200911180024-4186241/srimulyani-bicara-dampak-psbb-luar-biasa-serius.
- Lidwina Andrea *Penurunan Penghasilan dan Kebutuhan Bantu- an Sosial Responden di Surabaya akibat Covid-19*, 21 Juli 2020,
  https://databoks.katadata.co.id/datapublish/2020/07/21/bagaimana-kondisiekonomi-warga-surabaya-akibat-pandemi-covid-19.
- Andira, Covid-19: Bansos Terdampak Corona Di Surabaya Tidak Tepat Sasaran, 08

  Mei 2020, https://www.suarakarya.id/detail/110903/Bansos-Terdampak-CoronaDi-Surabaya-Tidak-Tepat-Sasaran
- Tim Nasional Percepatan Kemiskinan *Wapres Pimpin Rapat TNP2K Bahas Target Penurunan Kemiskinan dan Stunting 2024*, 11 Februari 2020. http://WWW.tnp2k.go.id.articles/wapres-pimpin-rapat-tnp2k-bahas-target-penurunan-kemiskinan-dan-stunting-2024.

Shihab, M.Q., 2005, *Tafsir Al-Misbah*, Edisi Empat. Jakarta: Lentera Hati.

Utami, D.R., 2018, Aplikasi Monitoring Keluarga Miskin Menggunakan Metode

- K-Means Clustering Berbasis Moblile GIS (Studi Kasus: PKH Kec Kedung Kendang Kota Malang). Malang: Universitas Islam Negeri Malang.
- Febianto, N.I. dan Nico, D.P., 2019, Analisis Clustering K-Means Pada Data Informasi Kemiskinan di Jawa Barat Tahun 2018. Jurnal SISFOKOM.
- Kallestal, C., Elmer, Z.B., Wilton, P., Oleg, S., Katarina, E.S., Mariela, C., Lars, A.P. and Rodolfo P., 2020, Assessing the Multiple Dimensions of Poverty. Data Mining Approaches to the 2004–14 Health and Demographic Surveillance System in Cuatro Santos, Nicaragua. Frontiers in Public Health.
- Khairatia A.F., Adlinaa, A.A., Hertonoa, .G.F. dan Handaria, B.D., 2019, *Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan K-Means MMCA*. Jurnal Prosiding Seminar Nasional Matematika.
- Harini, N.H., Cahyo, P. dan Fikri, A.N., 2020, Segmentasi pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python. Jurnal JAMIKA.
- Isdijoso, Widjajanti, Asep, S., dan Akhmadi, 2016, Penetapan Kriteria Dan Variabel Pendataan Penduduk Miskin Yang Komprehensif Dalam Rangka Perlindungan Penduduk Miskin Di Kabupaten/Kota. The SMERU Research Institute.
- Ulya, H.N., 2018, *Paradigma Kemiskinan dalam Perspektif Islam dan Konvensio-nal*, Journal of Islamic Economics and Business.
- Duraesa, M.A., 2016, *Kemiskinan dalam Al-Quran: Suatu Tinjauan Teologis*, Ponorogo: Uwais Inspirasi Indonesia.
- Jiawei Han, Micheline, K., and Jian, P., 2011, *Data Mining Concepts and Techniques*, 3rd ed, USA: an imprint of Elsevier.

- Prasetyo E., 2014, *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: ANDI.
- Wanto A., Muhammad, N.H.S., Agus, P.W., Dedy, H., Ni Luh, W., Sri, R.G., Darmawan, N., Edi, S.N, Muhammad, R.L., Sarini, V.D., dan Cahyo, P., 2020, *Data Mining Algoritma dan Impementasi*, 1st ed., United States of America.
- Tan P., Michael, S., and Vipin, K., 2006. *Introduction to Data Mining*, 1st ed., United States of America.
- Purnama, B., 2019, Pengantar Mechine Learning. Bandung: Informatika.
- Fhadli, M., 2020, *Data Mining Dengan Python Untuk Pemula*. Ternate: Guepedia.
- Harani, N.H., Cahyo, dan P., Fikri, A.N., 2020, Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Al-Goritma K-Means Berbasis Python, Jurnal Manajemen Informatika.
- Bisri, A., 2020, Terjemah Alfaraidul Bahiyyah, Rembang: Menara Kudus.