

**IMPLEMENTATION OF THE K-MEANS CLUSTERING ALGORITHM IN  
ANALYZING PUBLIC SATISFACTION REGARDING PUBLIC SERVICES (STUDI  
CASE: BALAI PENGUJIAN STANDAR INSTRUMEN TANAMAN INDUSTRI DAN  
PENYEGAR)**

**Atika Juhaedah Alifah<sup>\*1</sup>, Sudin Saepudin<sup>2</sup>, Carti Irawan<sup>3</sup>**

<sup>1,2,3</sup>Information System, Faculty of Computer Engineering and Design, Universitas Nusa Putra, Indonesia  
Email: [atika.j.alifah@gmail.com](mailto:atika.j.alifah@gmail.com), [sudin.saepudin@nusaputra.ac.id](mailto:sudin.saepudin@nusaputra.ac.id), [carti.irawan@nusaputra.ac.id](mailto:carti.irawan@nusaputra.ac.id)

(Article received: May 23, 2024; Revision: June 25, 2024; published: August 11, 2024)

**Abstract**

With the development of today's modern era, publik service is an important and very necessary thing because it is one of the benchmarks for seeing publik trust and satisfaction with the services provided by an agency. One of the agencies that carries out publi services is the Balai Pengujian Standar Instrumen Tanaman Industri dan Penyegar (BPSI TRI), a government agency under the Ministry of Agriculture. There are a lot of people who will receive services in 2023. Therefore, publik service officers find it difficult to determine publik satisfaction in order to optimize the services provided. To determine community satisfaction, data mining calculations were carried out using the K-Means clustering algorithm method with Community Satisfaction Index (IKM) data in 2023 using 3 (three) categories including unsatisfactory (C1), satisfactory (C2) and very satisfactory) and 2 attributes, namely the behavior of service officers (U7) as well as handling complaints, suggestions and input (U8) then carried out calculations using Microsoft Excel and got the results that C1 (unsatisfactory) 14 respondents, C2 (satisfactory) 39 respondents and C3 (very satisfactory) 98 respondents. Meanwhile, from the results of calculations using python testing, the results showed that C1 (unsatisfactory) was 9 respondents, C2 (satisfactory) was 39 respondents and C3 (very satisfactory) was 103 respondents.

**Keywords:** K-Means Clustering, Microsoft Excel, Publik Service, Python.

**IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM  
MENGANALISIS KEPUASAN MASYARAKAT MENGENAI PELAYANAN PUBLIK  
(STUDI KASUS: BALAI PENGUJIAN STANDAR INSTRUMEN TANAMAN  
INDUSTRI DAN PENYEGAR)**

**Abstrak**

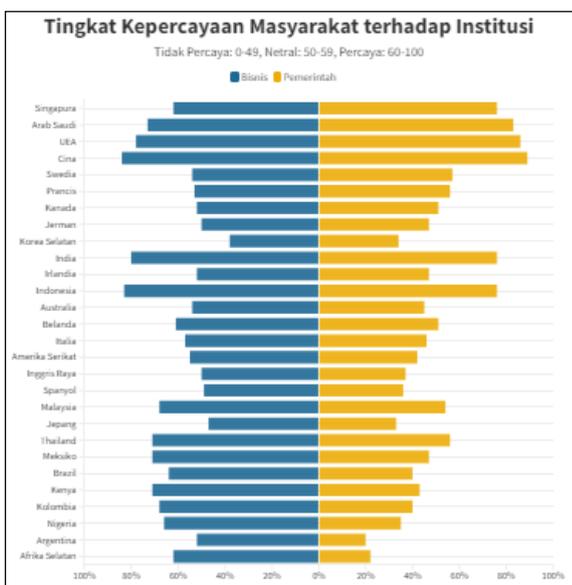
Dengan perkembangan zaman yang modern saat ini, pelayanan publik merupakan salah satu hal penting dan sangat diperlukan karena menjadi salah satu tolak ukur untuk melihat kepercayaan dan kepuasan masyarakat terhadap pelayanan yang diberikan oleh sebuah instansi. Salah satu instansi yang melaksanakan pelayanan publik yaitu Balai Pengujian Standar Instrumen Tanaman Industri dan Penyegar (BPSI TRI) salah satu instansi pemerintahan di bawah kementerian pertanian. Pada tahun 2023 banyak masyarakat yang menerima pelayanan oleh karena itu, petugas pelayanan publik kesulitan untuk mengetahui kepuasan masyarakat guna mengoptimalkan pelayanan yang diberikan. Untuk mengetahui kepuasan masyarakat maka dilakukan perhitungan data *mining* dengan metode algoritma *K-Means clustering* dengan data Indeks Kepuasan masyarakat (IKM) pada tahun 2023 menggunakan 3 (Tiga) kategori diantaranya kurang memuaskan (C1), memuaskan (C2) serta sangat memuaskan) dan 2 atribut yaitu perilaku petugas pelayanan (U7) serta penanganan pengaduan, saran dan masukan (U8) kemudian melakukan implementasi perhitungan menggunakan Microsoft Excel dan mendapatkan hasil bahwa C1 (kurang memuaskan) 14 responden, C2 (memuaskan) 39 responden dan C3 (sangat memuaskan) 98 responden. sedangkan, dari hasil perhitungan menggunakan pengujian python mendapatkan hasil bahwa C1 (kurang memuaskan) 9 responden, C2 (memuaskan) 39 responden dan C3 (sangat memuaskan) 103 responden.

**Kata kunci:** K-Means Clustering, Microsoft Excel, Pelayanan Publik, Python.

## 1. PENDAHULUAN

### 1.1. Latar Belakang

Dengan perkembangan zaman yang modern saat ini, pelayanan publik merupakan salah satu yang diperlukan pada sebuah instansi dan merupakan hal penting untuk menjadi tolak ukur untuk melihat kepercayaan dan kepuasan masyarakat terhadap pelayanan yang diberikan. Pada survei oleh *Edelman Trust Barometer* yang dikutip dari data tempoyang menyebarkan kuesioner di 28 negara menunjukkan bahwa tingkat kepercayaan masyarakat terhadap institusi lebih rendah dibandingkan tingkat kepercayaan terhadap bisnis. Salah satunya Negara Indonesia yang berada di urutan ke-12 yang masyarakatnya 83% lebih percaya terhadap institusi bisnis dibandingkan dengan kepercayaan terhadap institusi pemerintahan dengan persentase 76%. Seperti yang tersaji pada grafik berikut: [1]

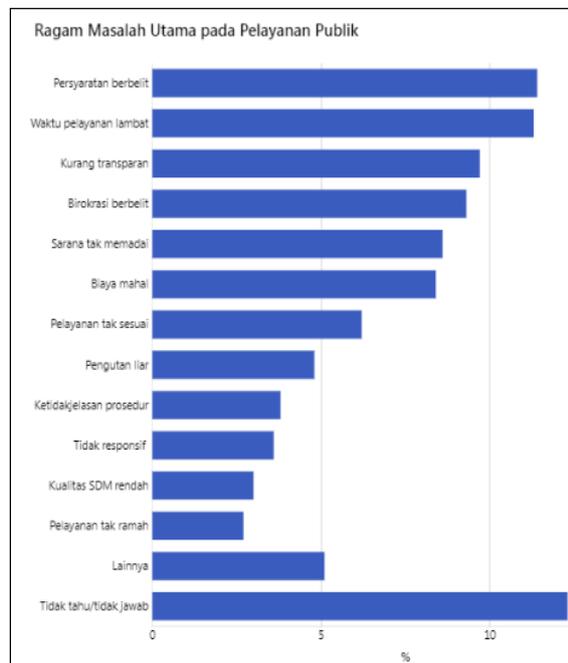


Gambar 1. Grafik Tingkat Kepercayaan masyarakat Terhadap Institusi

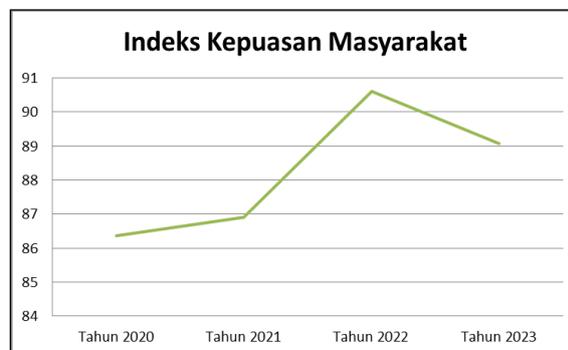
Selanjutnya survei yang dilakukan oleh Populi Center pada tanggal 1-9 desember 2021 yang dikutip dari katadata mengenai permasalahan yang terjadi di Indonesia pada pelayanan publik terdapat 11,4% masyarakat mengeluhkan persyaratan yang berbelit, 8,6% , mengeluhkan sarana dan prasarana yang tidak memadai, 2,7% mengeluhkan pelayanan yang kurang ramah serta pelayanan lainnya[2], dapat dilihat pada gambar 2.

Dengan demikian, kualitas suatu instansi tergantung pada pelayanan yang diberikan serta menjadi salah satu indicator utama untuk mengukur Tingkat kepuasan masyarakat terhadap kinerja suatu instasi pemerintah dan lembaha publik lainnya. Salah satu instansi yang menyelenggarakan pelayanan publik yaitu Balai Pengujian Standar Instrumen Tanaman Industri dan Penyegar (BPSI TRI) sesuai dengan peraturan pemerintah yang mewajibkan seluruh instansi menyelenggarakan pelayanan publik.

Pada setiap kegiatan pelayanan yang dilakukan oleh BPSI TRI petugas pelayanan akan memberikan formulir mengenai survei kepuasan masyarakat terhadap pelayanan yang diberikan. Sehingga, mendapatkan grafik Indeks Kepuasan masyarakat (IKM) pada gambar 3.



Gambar 2. Grafik Ragam Masalah Pada Pelayanan Publik



Gambar 3. Grafik IKM BPSI TRI Tahun 2020-2023

Pada hasil grafik IKM pelayanan yang diberikan oleh BPSI TRI tergolong sangat baik dan menunjukkan bahwa IKM pada tahun 2020 bernilai 86.37, tahun 2021 bernilai 96.91 dan mengalami penurunan di tahun 2023 yang bernilai 89.90.[3] kenaikan dan penurunan tersebut terjadi karena jumlah responden tidak selalu sama pada setiap tahunnya dan ini menyulitkan petugas pelayanan publik untuk mengetahui kepuasan masyarakat terhadap pelayanan yang diberikan.

Dengan adanya permasalahan tersebut maka diperlukan suatu *clustering* atau pengelompokkan untuk dapat melihat kepuasan masyarakat terhadap pelayanan yang diberikan. Hal ini dilakukan untuk mengoptimalkan dan upaya meningkatkan kualitas pelayanan publik di BPSI TRI.

Terdapat penelitian sebelumnya yang sejalan dengan penelitian ini dengan judul ‘Mengukur Tingkat Kepuasan masyarakat Pada Pelayanan Kepolisian Resor (POLRES) Dumai Menggunakan Algoritma K-Means Clustering’ yang menggunakan 79 data responden kemudian mendapatkan hasil 67 responden merasa sangat puas dan 12 responden cukup puas terhadap pelayanan yang diberikan.[4]

Penelitian selanjutnya dengan judul ‘Penerapan Data Mining *K-Means Clustering* untuk Mengelompokkan Berbagai Jenis Merk Mesin Cuci’. Dengan 25 data merk mesin cuci dan mendapatkan Hasil yang menyatakan bahwa pada cluster 1 berjumlah 5 data merk mesin cuci, cluster 2 berjumlah 8 data merk mesin cuci, dan cluster 3 berjumlah 12 merk mesin cuci. Penelitian ini dijadikan acuan untuk memilih merk mesin cuci yang tepat sesuai dengan kebutuhan dan juga keinginan.[5] Penelitian yang berjudul ‘Implementasi Metode K-Means Dalam Pengklasteran Daerah Pungutan Liar Di Kabupaten Sukabumi (Studi Kasus: Dinas Kependudukan dan Pencatatan Sipil)’ yang menghasilkan bahwa C1 (pungutan liar tinggi) terdapat 13 kecamatan, C2 (pungutan liar sedang) terdapat 19 kecamatan, dan C3 (pungutan liar rendah) terdapat 16 kecamatan.[6] Dan penelitian selanjutnya dengan judul ‘Analysis Of Stunting Spread In Mentawai Island Regency Using K-Means Clustering Method’ yang membahas tentang permasalahan kritis Kesehatan dan gizi anak di Kabupaten Mentawai, dan dengan menggunakan 10 data kabupaten yang berada di Mentawai dan menghasilkan bahwa C1 terdapat 6 kabupaten, C2 berjumlah 2 kabupaten, dan C3 terdapat 2 kabupaten.[7]

Perbedaan dari penelitian sebelumnya yang sudah diuraikan yaitu implementasi algoritma k-means menggunakan Microsoft excel dan Bahasa pemrograman python. Penelitian ini bertujuan untuk mengetahui kepuasan masyarakat terhadap pelayanan yang dilakukan oleh BPSI TRI, kemudian hasil clustering ini dapat menjadi acuan untuk meningkatkan serta mengevaluasi kinerja petugas pelayanan publik untuk memberikan pelayanan yang lebih maksimal.

## 1.2. Kajian Teori

### 1.2.1. Pelayanan Publik

Merupakan sebuah kegiatan atau serangkaian kegiatan dalam rangka pemenuhan dengan memberikan pelayanan atau melayani kebutuhan masyarakat, sesuai dengan tata cara dan aturan yang ditentukan dan bertujuan untuk memberikan kepuasan terhadap penerima layanan. [8].

### 1.2.2. Indeks Kepuasan masyarakat

Merupakan pengukuran perbedaan antara yang ingin ditawarkan oleh sebuah instansi dan diwujudkan oleh masyarakat dengan memberikan suatu produk atau jasa serta memberikan alasan

penggunaan IKM apapun bentuk organisasinya. Keberhasilan suatu instansi terlihat dari program mutu tersebut dapat diukur dari kepuasan masyarakat berdasarkan pendapat yang mengatakan bahwa untuk mengetahui mutu instansi.[9].

### 1.2.3. Kepuasan masyarakat

Merupakan perasaan senang atau tidaknya masyarakat yang berasal dari perbandingan antara kesenangan terhadap aktivitas dan suatu layanan produk dengan harapannya yaitu kepuasan. [10].

### 1.2.4. Data Mining

Merupakan suatu proses menemukan pola dengan memeriksa dalam sekumpulan data yang besar yang tersimpan dalam penyimpanan dengan menggunakan teknik dan pola statistik dan matematika yang menghasilkan pola baru.[11].

### 1.2.5. Algoritma K-Means

Merupakan salah satu metode clustering non-hirarki yang mengelompokkan data dalam bentuk satu cluster atau lebih yang memiliki karakteristik yang sama [12].

### 1.2.6. Clustering

Merupakan sebuah proses algoritma dalam mengelompokkan sejumlah data tertentu dengan sedemikian rupa sehingga objek yang tidak serupa menjadi anggota cluster, dalam setiap cluster memiliki data yang memiliki kemiripan dihitung dengan jarak.[13].

### 1.2.7. Python

Merupakan salah satu bahasa pemrograman tingkat tinggi karena perintah atau kode program yang dipakai sudah mirip dengan bahasa manusia.[14].

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif yang merupakan pendekatan yang menggambarkan atau menjelaskan suatu masalah yang dapat digeneralisasikan karena berdasarkan kenyataan sebenarnya bahwa penelitian berhubungan dan melibatkan data. Metode penelitian yang digunakan adalah algoritma K-means, algoritma K-means ini melakukan clusterisasi pada data yang akan diuji. Dengan *flowchart* pada gambar 4.

Tahap pengelompokkan menggunakan algoritma K-Means menurut M. Syukron Nawawi adalah sebagai berikut: [15]

1. Jumlah k (cluster) dan centroid awal secara acak
2. Jarak data dan pusat cluster dihitung menggunakan *Euclidean Distance* yang dirumuskan sebagai berikut:

$$D(i, j) = \sqrt{(X1i - X1j)^2} + \sqrt{(X2i - X2j)^2} + \dots + \sqrt{(Xki - Xkj)^2} \quad (1)$$

Keterangan :

$D(i, j)$  = Jarak data ke pusat *Cluster*

$Xki$  = Data ke *i* pada atribut data ke *k*

$Xkj$  = Titik pusat ke *j* pada atribut ke *k*

3. Mengelompokkan data ke dalam cluster dengan jarak yang paling dekat.
4. Menghitung pusat cluster yang baru dengan mencari nilai rata-rata pada kelompok data dengan rumus :

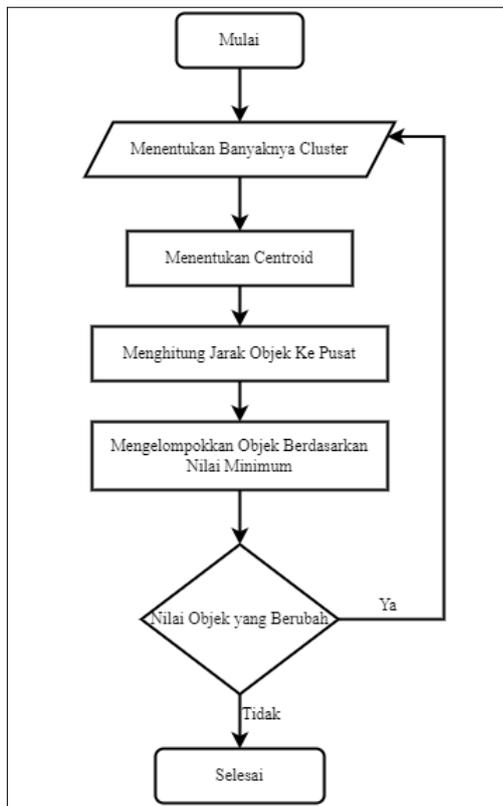
$$\mu_j (t + 1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j \quad (2)$$

Keterangan:

$\mu_j (t + 1)$  = Nilai centroid baru pada iterasi

$\frac{1}{N_{sj}} \sum_{j \in s_j} x_j$  = Nilai centroid baru pada iterasi

5. Pusat cluster akan ditentukan bila semua data telah ditetapkan dalam cluster terdekat.
6. Proses penentuan pusat cluster dan penempatan data dalam cluster diulang hingga nilai centroid tidak berubah lagi.



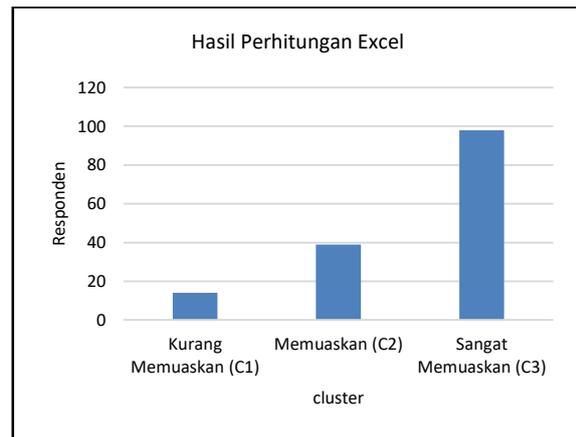
Gambar 4. Flowchart Algoritma K-Means

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Hasil

Data sampel yang digunakan yaitu data Survei Kepuasan masyarakat (SKM) Tahun 2023 dengan

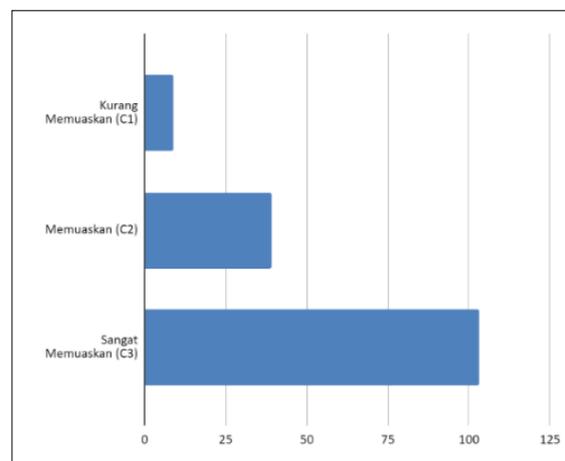
banyak jumlah data yaitu 151 data dengan 9 atribut yaitu persyaratan (U1), Sistem, Mekanisme dan Prosedur (U2), waktu penyelesaian (U3), biaya/tarif (U4), prodeuk spesifikasi jenis pelayanan (U5), kompetensi pelaksana (U6), perilaku pelaksana pelayanan (U7) dan penanganan pengaduan, saran dan masukan (U8) serta sarana dan prasarana (U9). Dari data yang ada, dihitung menggunakan Microsoft Excel dan menghasilkan data seperti grafik berikut:



Gambar 5. Hasil Perhitungan Excel

Dari grafik diatas dapat disimpulkan bahwa perhitungan dari Microsoft Excel menghasilkan C1 (kurang memuaskan) terdapat 14 responden, C2 (memuaskan) terdapat 39 responden dan C3 (sangat memuaskan) terdapat 98 responden.

Sedangkan pada pehitungan atau pengujian menggunakan python mendapatkan hasil bahwa C1 (kurang memuaskan) terdapat 9 reponden, C2 (memuaskan) terdapat 39 responden dan C3 (sangat memuaskan) terdapat 103 responden. Seperti grafik berikut:



Gambar 6. Hasil Perhitungan python

#### 3.1.1. Processing Data

Setelah mendapatkan data SKM BPSI TRI pada tahun 2023 dilakukan *preprocessing* sebelum melakukan implementasi dan perhitungan. Berikut adalah tahapan *preprocessing* data yang dilakukan:

1. Seleksi Atribut

Dari data awal yang didapatkan terdapat data yang sangat beragam dan memiliki beberapa atribut yang diperlukan pada pelayanan publik. Namun, setelah dilakukan penyeleksian atribut maka peneliti hanya menggunakan (U7) dan (U8) dikarenakan dua atribut tersebut sangat penting untuk keberlangsungan suatu pelayanan.

2. Data Cleaning

Setelah menentukan atribut, selanjutnya dilakukan pembersihan data yang tidak sesuai sebelum dilakukan proses pengolahan data lebih lanjut. Berikut adalah hasil dari seleksi atribut:

Tabel 1. Hasil Data Cleaning

No	Responden	Atribut	
		U7	U8
1	Pelajar/Mahasiswa	4	4
2	Pelajar/Mahasiswa	4	4
3	Pelajar/Mahasiswa	4	2
4	Pelajar/Mahasiswa	4	4
...	...	...	...
149	Pelajar/Mahasiswa	4	4
150	Pelajar/Mahasiswa	4	4
151	Pelajar/Mahasiswa	3	4

3. Transformasi Data

Dari hasil proses seleksi atribut dan cleaning data sebelumnya, proses yang dilakukan selanjutnya adalah melakukan perubahan data yang memiliki data uang tidak dapat diolah secara matematis menjadi data yang dapat diolah. Transformasi ini berupa data teks menjadi data angka (numerik).

3.1.2. Implementasi K-Means Clustering Menggunakan Microsoft Excel

Setelah melakukan preprocessing data. Tahapan selanjutnya yaitu pengimplementasian metode K-Means menggunakan Microsoft Excel dengan beberap tahapan yang dilakukan, seperti berikut:

1. Menentukan nilai k (cluster) dan centroid secara acak, seperti berikut:
2. Menghitung cluster dengan kategori yang telah ditentukan yaitu C1 (kurang memuaskan), C2 (memuaskan) dan C3 (sangat memuaskan) dengan rumus *Euclidean Distance* sebagai berikut.

Tabel 3. Hasil Perhitungan Jarak Terdekat

No	Responden	Atribut		C1	C2	C3	Jrk Terdkt
		U7	U8				
1	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0	1,0
2	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0	1,0
3	Pelajar/Mahasiswa	4	2	1,4	2,2	2,2	1,4
4	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0	1,0
...	...	...	...	...	...	...	...
149	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0	1,0
150	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0	1,0
151	Pelajar/Mahasiswa	3	4	1,0	0,0	0,0	0,0

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid ditentukan dengan nilai jarak terkecil, sehingga mendapatkan hasil sebagai berikut:

$$D(i, j) = \sqrt{(X1i - X1j)^2} + \sqrt{(X2i - X2j)^2} + \dots + \sqrt{(Xki - Xkj)^2} \quad (1)$$

Sehingga pada iterasi 1 (Satu) mendapatkan hasil perhitungan sebagai berikut:

C1 = (3,3)

$$[1] D(i, j) = \sqrt{(4 - 3)^2 + (3 - 4)^2} = 1,4$$

$$[2] D(i, j) = \sqrt{(4 - 3)^2 + (3 - 4)^2} = 1,4$$

$$[3] D(i, j) = \sqrt{(4 - 3)^2 + (3 - 2)^2} = 1,4$$

[4] ... ..

C2 = (3,4)

$$[1] D(i, j) = \sqrt{(4 - 3)^2 + (4 - 4)^2} = 1,0$$

$$[2] D(i, j) = \sqrt{(4 - 3)^2 + (4 - 4)^2} = 1,0$$

$$[3] D(i, j) = \sqrt{(4 - 3)^2 + (2 - 4)^2} = 1,0$$

[4] ... ..

C3 = (3,4)

$$[1] D(i, j) = \sqrt{(4 - 3)^2 + (4 - 4)^2} = 1,0$$

$$[2] D(i, j) = \sqrt{(4 - 3)^2 + (4 - 4)^2} = 1,0$$

$$[3] D(i, j) = \sqrt{(4 - 3)^2 + (4 - 2)^2} = 1,0$$

[4] ... ..

Perhitungan tersebut dilanjutkan hingga data terakhir sehingga mendapatkan hasil seperti tabel di bawah ini:

Tabel 2 Hasil Perhitungan cluster

No	Responden	Atribut		C1	C2	C3
		U7	U8			
1	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0
2	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0
3	Pelajar/Mahasiswa	4	2	1,4	2,2	2,2
4	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0
...	...	...	...	...	...	...
149	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0
150	Pelajar/Mahasiswa	4	4	1,4	1,0	1,0
151	Pelajar/Mahasiswa	3	4	1,0	0,0	0,0

3. Setelah melakukan perhitungan cluster, selanjutnya melakukan perhitungan jarak terdekat dengan menghitung nilai rata-rata paling minimal sehingga mendapatkan hasil sebagai berikut:

Tabel 4. Hasil Clusterisasi

No	Responden	Atribut		Jarak Terdkt	Klpk
		U7	U8		
1	Pelajar/Mahasiswa	4	4	1,0	3
2	Pelajar/Mahasiswa	4	4	1,0	3
3	Pelajar/Mahasiswa	4	2	1,4	1

4	Pelajar/Mahasiswa	4	4	1,0	3
...	...	..	...	1,0	3
149	Pelajar/Mahasiswa	4	4	1,0	3
150	Pelajar/Mahasiswa	4	4	1,0	3
151	Pelajar/Mahasiswa	3	4	0,0	2

Dari perhitungan yang didapatkan pada table di atas, kemudian disimpulkan sesuai dengan masing-masing cluster, sehingga pada iterasi 1 dapat disimpulkan bahwa C1 (kurang memuaskan) 14 responden, C2 (memuaskan) 39 responden dan C3 (sangat memuaskan) 98 responden. seperti pada table berikut:

Tabel 5. Kesimpulan Iterasi 1

Kesimpulan Iterasi 1	
Cluster 1	14
Cluster 2	39
Cluster 3	98
Total	151

- Memperbaharui nilai centroid baru yang diperoleh dengan rumus yang sudah ditetapkan, kemudian akan menghasilkan centroid baru yang akan dihitung pada iterasi selanjutnya.
- Melakukan perhitungan ulang dari iterasi 2 hingga seterusnya dengan nilai centroid baru hingga anggota tiap cluster memiliki jumlah yang sama. Dengan melakukan beberapa kali perhitungan iterasi terdapat hasil kesimpulan dari perhitungan iterasi 2, seperti berikut:

Tabel 6. Kesimpulan Iterasi 2

Kesimpulan Iterasi 2	
Cluster 1	13
Cluster 2	39
Cluster 3	99
Total	151

Pada hasil perhitungan seperti table diatas dapat disimpulkan bahwa C1 (kurang memuaskan) 13 responden, C2 (memuaskan) 39 responden dan C3 (sangat memuaskan) 99 responden. Karena mengalami perbedaan jumlah nilai pada masing-masing cluster maka perhitungan dilanjutkan dengan iterasi 3, dan mendapatkan kesimpulan seperti tabel berikut:

Tabel 7. Kesimpulan Iterasi 3

Kesimpulan Iterasi 3	
Cluster 1	15
Cluster 2	38
Cluster 3	98
Total	151

Pada iterasi 3 dapat disimpulkan bahwa C1 (kurang memuaskan) 15 responden, C2 (memuaskan) 38 responden dan C3 (sangat memuaskan) 98 responden. Perhitungan masih dilanjutkan ke iterasi 4 karena iterasi 3 masih menghasilkan jumlah cluster yang berbeda dan pada iterasi 4 menghasilkan seperti tabel berikut:

Tabel 8. Kesimpulan Iterasi 4

Kesimpulan Iterasi 4	
Cluster 1	14
Cluster 2	39
Cluster 3	98
Total	151

Perhitungan pada iterasi 4 mendapatkan kesimpulan bahwa C1 terdapat 14 responden, C2 terdapat 39 responden dan C3 terdapat 98 responden. Pada perhitungan iterasi 4 ini, masih mengalami perubahan jumlah nilai maka dilanjutkan dengan perhitungan iterasi 5 dan mendapatkan hasil sebagai berikut:

Tabel 9. Hasil Perhitungan Iterasi 5

No	Responden	Atribut		Klpk	No awal
		U7	U8		
1	Pelajar/Mahasiswa	4	2	1	3
2	Pelajar/Mahasiswa	3	3	1	7
3	Pelajar/Mahasiswa	3	2	1	23
4	Pelajar/Mahasiswa	4	3	1	34
...	...	...	...	...	...
149	Pelajar/Mahasiswa	4	4	3	148
150	Pelajar/Mahasiswa	4	4	3	149
151	Pelajar/Mahasiswa	4	4	3	150

Karena perhitungan iterasi 4 seperti di atas, bahwa nilai iterasi 5 sama dengan iterasi 4 maka perhitungan iterasi ini berakhir pada iterasi 5 dengan hasil kesimpulan sebagai berikut:

Tabel 10. Kesimpulan Iterasi 5

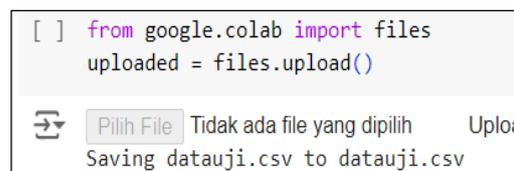
Kesimpulan Iterasi 5	
Cluster 1	14
Cluster 2	39
Cluster 3	98
Total	151

Pada perhitungan di atas, dapat disimpulkan bahwa iterasi 5 mendapatkan hasil C1 terdapat 14 responden, C2 terdapat 39 responden dan C3 terdapat 98 responden.

### 3.1.3. Implementasi K-Means Clustering Menggunakan Python

Berdasarkan hasil data di proses menggunakan Microsoft Excel yang sudah dijelaskan sebelumnya, maka tahapan selanjutnya yaitu implementasi menggunakan python dengan proses metode K-means Clustering sebagai berikut:

- Melakukan upload file menggunakan data yang sudah dilakukan Praproses, seperti berikut:



- Melakukan import library sesuai dengan kebutuhan proses pengujian system. Seperti gambar dibawah:

```
#importLibrary
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import random as rd
from collections import defaultdict
import matplotlib.cm as cm
```

Gambar 8. Proses Import Library

Library yang digunakan adalah Numpy yang akan digunakan untuk memproses atau mengkonversi data kedalam bentuk array, Matplotlib digunakan untuk memvisualisasikan data agar mudah untuk diterjemahkan, pandas digunakan untuk proses manipulasi tabel dalam bentuk numerik, dan Scikit Learn merupakan library yang berperan sangat penting untuk sebuah proses implementasi algoritma yang digunakan pada penelitian ini yaitu *K-Means Clustering*. [16]

- Menampilkan 5 (Lima) baris pertama dari data yang sudah di upload, seperti gambar berikut:

```
# Menampilkan 5 Baris Pertama
dataset = pd.read_csv('datauji.csv')
dataset.head(5)
```

Responden	Perilaku pelaksana pelayanan	Penanganan pengaduan, saran dan masukan
0 Pelajar/Mahasiswa	4	4
1 Pelajar/Mahasiswa	4	4
2 Pelajar/Mahasiswa	4	2
3 Pelajar/Mahasiswa	4	4
4 Pelajar/Mahasiswa	3	4

Gambar 9. Proses Menampilkan Data

- Tahapan selanjutnya, memasukkan data tersebut ke dalam variable dan menampilkan hasil berbentuk array. Seperti di bawah ini:

```
#Memasukan data set ke variabel dan menampilkan hasilnya
X = dataset.iloc[:, [1,2]].values
print(X)
```

```
[[4 4]
 [4 4]
 [4 2]
 [4 4]
 [3 4]
 [3 4]
 [3 3]
 [4 4]
 [4 4]
 [4 4]
 [4 4]]
```

Gambar 10. Proses Memasukkan ke dalam variable

- Menentukan cluster dan titik centroid secara acak dengan jumlah data 151 serta 3 kategori yaitu kurang memuaskan, memuaskan dan sangat memuaskan. Dengan melakukan konversi data ke dalam bentuk array dengan bantuan library Numpy dan melakukan reshape agar data tersebut dapat diolah.
- Berdasarkan ketentuan system yang berlaku, seperti gambar berikut:

```
[ ] #Menentukan jumlah cluster dan titik centroid
K = 3
m = 151
Centroids = np.array([]).reshape(2,0)

#STEP 1 : Menentukan letak centroid secara acak
for i in range(K):
    rand = rd.randint(0,m-1)
    Centroids = np.c_[Centroids,X[rand]]
```

Gambar 11. Proses Penentuan cluster dan centroid

- Melakukan implementasi algoritma K-Means Clustering dibantu dengan library yang tersedia pada mesin untuk menghitung iterasi, seperti berikut:

```
#step2.a Menyesuaikan koordinat centroid dengan euclidean distance
EuclideanDistance = np.array([]).reshape(m,0)
for k in range(K):
    tempJarak=np.sum((X-Centroids[:,k])**2,axis=1)
    EuclideanDistance=np.c_[EuclideanDistance,tempJarak]

C=np.argmax(EuclideanDistance,axis=1)+1
```

Gambar 12. Proses implementasi K-means

- Melakukan perhitungan jumlah iterasi dengan menyesuaikan titik koordinat pada centroid dengan rumus *Euclidean Distance* dengan menentukan jarak dari setiap cluster dan centroid yang ada, seperti gambar d bawah ini:

```
#step2.a Menyesuaikan koordinat centroid dengan euclidean distance
EuclideanDistance = np.array([]).reshape(m,0)
for k in range(K):
    tempJarak=np.sum((X-Centroids[:,k])**2,axis=1)
    EuclideanDistance=np.c_[EuclideanDistance,tempJarak]

C=np.argmax(EuclideanDistance,axis=1)+1
```

Gambar 13. Proses menyesuaikan titik koordinat dengan rumus *Euclidean Distance*

- Tahapan selanjutnya yaitu mengcluster data yang ada ke dalam centroid dengan membuat array kosong untuk setiap cluster kemudian memasukkannya ke dalam cluster, seperti gambar berikut:

```
#step2.b Mengclusterkan data ke centroid
Y={}
#membuat array kosong untuk setiap cluster
for k in range (K):
    Y[k+1]=np.array([]).reshape(2,0)

#memasukan data ke dalam cluster
for i in range(m):
    Y[C[i]]=np.c_[Y[C[i]],X[i]]

#tranpose array
for k in range(K):
    Y[k+1] = Y[k+1].T
    cluster += [k]
```

Gambar 14. Proses cluster ke dalam centroid

- Memasukkan nilai centroid akhir ke variable clustering kemudian akan menampilkan hasil dari centroid akhir. Proses dapat dilihat pada gambar berikut:

```
[ ] #Menampilkan Data Hasil Perhitungan K-Means
Y
```

```
{1: array([[4., 4.],
          [4., 4.],
          [4., 4.],
          [4., 4.],
          [4., 4.],
          [4., 4.],
          [4., 4.]])
```

Gambar 15. Memasukkan nilai centroid akhir ke dalam variable

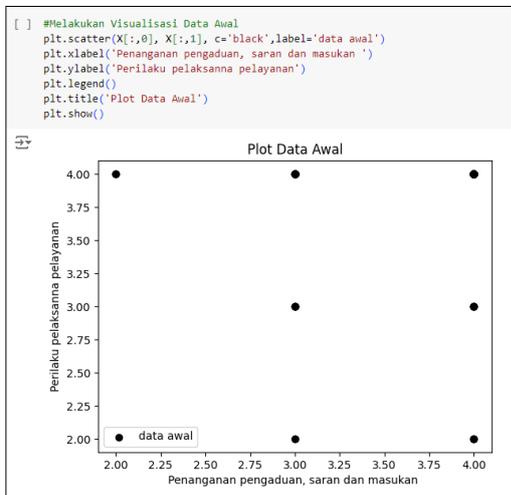
- Menampilkan hasil dari nilai akhir atau data hasil dari perhitungan K-Means seperti pada gambar berikut:

```
[ ] #Menampilkan Data Hasil Perhitungan K-Means
Y
{1: array([[4., 4.],
           [4., 4.],
           [4., 4.],
           [4., 4.],
           [4., 4.],
           [4., 4.],
           [4., 4.]])
```

Gambar 16. Proses menampilkan data hasil

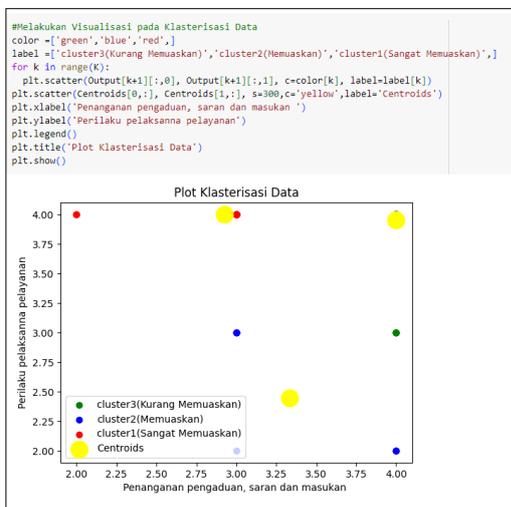
Data tersebut disajikan dalam bentuk angka yang dimulai dengan data responden pertama hingga responden akhir.

- Agar memudahkan dalam menerjemahkan data yang sudah diolah, maka dilakukan visualisasi dalam bentuk *scatterplot* dengan bantuan *library matplotlib* seperti pada gambar berikut berikut di bawah:



Gambar 17. Proses visualisasi dg scatterplot

- Dari hasil yang sudah didapatkan, kemudian dilakukan kembali visualisasi data menggunakan scatterplot agar mengetahui hasil clusterisasi dengan diberikan warna pada masing-masing cluster, seperti dibawah ini:



Gambar 18. Proses pewarnaan cluster

- Selanjutnya membuat tabel dengan menambahkan kolom cluster dari hasil perhitungan masing-masing cluster. Seperti di bawah ini:

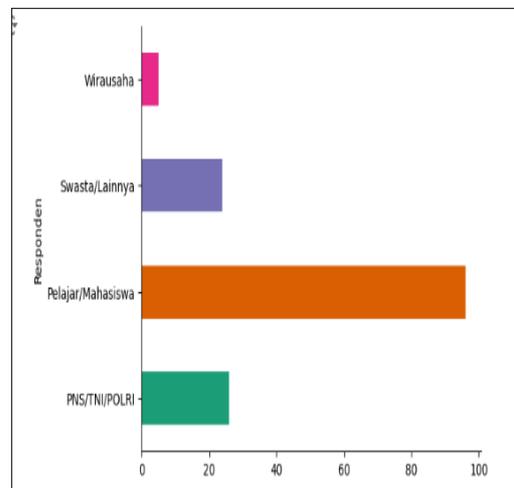
```
# Menambahkan kolom cluster pada dataset hasil visualisasi
dataset['Cluster']=c
dataset
```

Responden	Perilaku pelaksana pelayanannya	Penanganan pengaduan, saran dan masukan	Cluster
0 Pelajar/Mahasiswa	4	4	1
1 Pelajar/Mahasiswa	4	4	1
2 Pelajar/Mahasiswa	4	2	2
3 Pelajar/Mahasiswa	4	4	1
4 Pelajar/Mahasiswa	3	4	3
...	...	...	...
146 Pelajar/Mahasiswa	4	4	1
147 Pelajar/Mahasiswa	4	4	1
148 Pelajar/Mahasiswa	4	4	1
149 Pelajar/Mahasiswa	4	4	1
150 Pelajar/Mahasiswa	3	4	3

151 rows x 4 columns

Gambar 19. Proses penampilan data dengan tabel

- Tahapan selanjutnya menampilkan grafik pada masing-masing responden, sehingga dapat diketahui bahwa masyarakat yang menerima pelayanan BPSI TRI dari berbagai kalangan di dominasi oleh kalangan Pelajar/Mahasiswa. Seperti yang tersaji pada pada gambar 20.



Gambar 20. Grafik masing2 responden

- Tahapan terakhir yang dilakukan adalah menyimpan dan melakukan pengunduhan data hasil dari penggabungan dataset awal dengan cluster yang didapatkan. Sehingga data yang didapatkan akan mendapatkan informasi lebih lengkap dan bermanfaat. Berikut adalah proses yang dilakukan pada saat pengunduhan data.

```
[ ] #Menyimpan Hasil Pengolahan Data
dataset.to_csv('hasildatauji_ikmskripsi01.csv',index=False)
```

Gambar 21. Proses penyimpanan dan pengunduhan data

### 3.2. Pembahasan

Implementasi algoritma K-means untuk menentukan kepuasan masyarakat terhadap pelayanan yang diberikan oleh BPSI TRI dengan dikelompokkan dalam 3 kategori yaitu kurang memuaskan (C1), memuaskan (C2) dan sangat memuaskan (C3) seperti yang terdapat dalam hasil masing-masing perhitungan, sehingga menghasilkan perbedaan pada hasil data yang dapat diuraikan sebagai berikut:

1. Pada perhitungan yang dilakukan dengan Microsoft Excel dapat disimpulkan bahwa hasil dari perhitungan *K-Means* terdiri dari 5 iterasi dan menghasilkan bahwa 14 responden merasa kurang memuaskan (C1) terdiri dari 9 responden kalangan Pelajar/Mahasiswa, 3 responden dari kalangan PNS/TNI/POLRI dan 2 responden dari kalangan wirausaha; 39 responden merasa memuaskan (C2) terdiri dari 20 responden dari kalangan Pelajar/Mahasiswa, 10 responden dari kalangan PNS/TNI/POLRI, 7 responden dari kalangan Swasta/Lainnya serta 2 responden dari kalangan wirausaha; dan 98 responden merasa pelayanan yang diberikan sangat memuaskan (C3) terdiri dari 67 responden dari kalangan Pelajar/Mahasiswa, 13 responden dari kalangan PNS/TNI/POLRI, 17 responden dari kalangan Swasta/Lainnya dan 1 dari kalangan wirausaha. Sedangkan,
2. Pada perhitungan atau pengujian yang dilakukan dengan Bahasa Pemrograman python mendapatkan hasil bahwa 9 responden merasa kurang memuaskan (C1) terdiri dari 5 responden dari kalangan Pelajar/Mahasiswa, 3 responden dari kalangan PNS/TNI/POLRI serta 1 responden dari kalangan Wirausaha; 39 responden merasa memuaskan (C2) terdiri dari 20 responden dari kalangan Pelajar/Mahasiswa, 10 dari kalangan PNS/TNI/POLRI, 7 responden dari kalangan Swasta/Lainnya serta 2 responden lainnya dari kalangan wirausaha; dan masyarakat yang merasakan sangat memuaskan (C3) terhadap pelayanan terdapat 103 responden, yang diantaranya 71 responden dari kalangan Pelajar/Mahasiswa, 12 responden dari kalangan PNS/TNI/POLRI, 17 responden dari kalangan Swasta/Lainnya dan 2 responden lainnya dari kalangan Wirausaha.
3. Hasil dari perhitungan keduanya memiliki perbedaan karena memiliki cara implementasi yang berbeda dari pra-proses dan transformasi data, ketepatan floating point, kemampuan analisis dan visualisasi yang biasanya di python untuk memvisualisasikan menggunakan matplotlib.

### 4. DISKUSI

Berdasarkan pada penelitian ini menunjukkan bahwa pelayanan yang diberikan oleh BPSI TRI

tergolong sangat baik dikarenakan banyak masyarakat dari berbagai kalangan sangat puas terhadap pelayanan yang diberikan. Mengacu pada penelitian sebelumnya yang relevan dengan penelitian ini bahwa dengan menggunakan metode *Algoritma K-Means* dapat mengelompokkan kategori dengan cluster yang disediakan.

### 5. KESIMPULAN

Berdasarkan pada tahapan penelitian yang dilakukan mengenai analisis kepuasan masyarakat terhadap pelayanan publik yang dilakukan petugas pelayanan BPSI TRI, maka peneliti dapat menarik kesimpulan bahwa, penggunaan *K-Means Clustering* menggunakan beberapa tahapan sesuai dengan ketentuan metode *K-Means Clustering* itu sendiri diantaranya penentuan jumlah *k* (*Cluster*) dan *centroid* awal secara acak, menghitung jarak data pada pusat *Cluster* dengan menggunakan rumus *Euclidean Distance*, mengelompokkan data ke dalam *Cluster* dengan jarak paling dekat, menghitung pusat *Cluster* yang baru dengan mencari nilai rata-rata pada kelompok data, menentukan pusat *Cluster* ditentukan apabila seluruh data ditetapkan dalam *Cluster* terdekat, dan melakukan perhitungan ulang pada iterasi selanjutnya hingga jumlah *Cluster* tidak berubah lagi setelah menentukan pusat *Cluster* baru.

Pada penelitian ini berhasil mengimplementasikan algoritma *K-Means Clustering* menggunakan Microsoft Excel dan bahasa pemrograman python untuk menentukan kepuasan masyarakat terhadap pelayanan yang diberikan BPSI TRI dengan kategori kurang memuaskan (C1), memuaskan (C2) dan sangat memuaskan (C3).

Hasil clusterisasi terhadap pengujian data penelitian menggunakan Microsoft Excel berhasil diselesaikan pada iterasi 5 yang menyimpulkan bahwa C1 (kurang memuaskan) berjumlah 14 responden, C2 (memuaskan) berjumlah 39 responden, dan C3 (sangat memuaskan) berjumlah 98 responden.

Hasil clusterisasi terhadap pengujian python data penelitian menggunakan bahasa pemrograman python berhasil diselesaikan dengan hasil akhir bahwa C1 (kurang memuaskan) berjumlah 9 responden, C2 (memuaskan) berjumlah 39 responden, dan C3 (sangat memuaskan) berjumlah 103 responden.

Hasil dari perhitungan berbeda karena implementasi perhitungan keduanya memiliki perbedaan karena memiliki cara implementasi yang berbeda dari pra-proses dan transformasi data, ketepatan floating point, kemampuan analisis dan visualisasi yang biasanya di python untuk memvisualisasikan menggunakan matplotlib.

### DAFTAR PUSTAKA

- [1] Data Tempo and Faisal Javier, "Tingkat Kepercayaan masyarakat terhadap Institusi,"

- <https://data.tempo.co/data/1630/tingkat-kepercayaan-masyarakat-terhadap-institusi-bisnis-dan-pemerintah-siapa-yang-lebih-tinggi>.
- [2] Kata Data and Cindy Mutia Annur, "Ragam Masalah Utama pada Pelayanan Publik," <https://databoks.katadata.co.id/datapublish/2021/12/20/persyaratan-berbelit-keluhan-utama-masyarakat-terhadap-pelayanan-publik>.
- [3] Balai Pengujian Standar Instrumen Tanaman Industri, "Indeks Kepuasan masyarakat," <https://tanamanindustri.bsip.pertanian.go.id/informasi-publik/indeks-kepuasan-masyarakat>.
- [4] U. Burelia, G. Urva, A. Sellyana, S. Tinggi Teknologi Dumai, and D. Sekolah Tinggi Teknologi Dumai, "MENGUKUR TINGKAT KEPUASAN MASYARAKAT PADA PELAYANAN KEPOLISIAN RESOR(POLRES) DUMAI MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," 2022, doi: [10.52072/jutekinf.v10i1.354](https://doi.org/10.52072/jutekinf.v10i1.354)
- [5] W. A. Wahyuni and S. Saepudin, "PENERAPAN DATA MINING CLUSTERING UNTUK MENGELOMPOKKAN BERBAGAI JENIS MERK MESIN CUCI," 2021.
- [6] F. Sembiring, O. Octaviana, and S. Saepudin, "IMPLEMENTASI METODE K-MEANS DALAM PENGKLASTERAN DAERAH PUNGUTAN LIAR DI KABUPATEN SUKABUMI (STUDI KASUS: DINAS KEPENDUDUKAN DAN PENCATATAN SIPIL)," *Jurnal Tekno Insentif*, vol. 14, no. 1, pp. 40–47, Apr. 2020, doi: [10.36787/jti.v14i1.165](https://doi.org/10.36787/jti.v14i1.165).
- [7] A. K. Vadreas, G. N. Yuhandika, H. Fryonanda, M. I. Lubis, U. I. Arsyah, and D. W. S. Nirad, "Analysis of Stunting Spread in Mentawai Islands Regency Using K-Means Clustering Method," 2024, pp. 1015–1031. doi: [10.2991/978-94-6463-364-1\\_93](https://doi.org/10.2991/978-94-6463-364-1_93).
- [8] D. H. H. S. H. W. A. A. Muchtar, "Peran Media Komunikasi Pemerintahan dan Perilaku Birokrasi Dalam Pelayanan Publik (Studi Pada Kecamatan Tarogong Kaler garut)," *JURNAL KOMUNIKASI & ADMINISTRASI PUBLIK*, vol. 10, no. Jurnal Professional, pp. 179–188, Jun. 2023, doi: [10.37676/professional.v10i1.3932](https://doi.org/10.37676/professional.v10i1.3932)
- [9] A. Kepuasan masyarakat Terhadap Pelayanan Publik Berdasarkan Indeks Kepuasan masyarakat di Kantor Kecamatan Ayah Kabupaten Kebumen Indah Nur, aini Sulisty, S. Partiw, Ediwidjojo, and D. oleh Politeknik Dharma Patria Kebumen, "Jurnal E-Bis (Ekonomi-Bisnis)," vol. 4, no. 2, pp. 276–286, 2020, doi: [10.37339/jurnal](https://doi.org/10.37339/jurnal).
- [10] A. Budiono, H. Manurung, and S. Syahputra, "Penilaian Kinerja Pegawai Desa Menggunakan Algoritms K-Means Berdasarkan Index Kepuasan masyarakat (Kantor Desa Padang Brahrang)," *Agustus*, vol. 6, no. 3, 2022.
- [11] Haris Kurniawan, Sarjon Defit, and Sumijan, "Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal," *Journal of Applied Computer Science and Technology*, vol. 1, no. 2, pp. 80–89, Dec. 2020, doi: [10.52158/jacost.v1i2.102](https://doi.org/10.52158/jacost.v1i2.102).
- [12] R. M. Sagala, "Prediksi Kelulusan Mahasiswa Menggunakan Data mining Algoritma K-means." doi: [10.36342/teika.v11i2.2610](https://doi.org/10.36342/teika.v11i2.2610).
- [13] S. Oop Sofiyah and R. Dinar Dana, "ANALISIS EFEKTIVITAS PELAYANAN PUBLIK MENGGUNAKAN K-MEANS CLUSTERING DI KECAMATAN SUKAGUMIWANG," 2023, doi: [10.36040/jati.v7i2.6536](https://doi.org/10.36040/jati.v7i2.6536)
- [14] A. E. Falentino Sembiring, "Bahasa Ular untuk Pemrograman Python," 2020.
- [15] M. Syukron Nawawi, F. Sembiring, and A. Erfina, "Implementasi Algoritma K-Means Clustering Menggunakan Orange Untuk Penentuan Produk Busana Muslim Terlaris," *PROGRAM STUDI TEKNIK INFORMATIKA-UNIVERSITAS PGRI MADIUN*, pp. 789–797, 2021.
- [16] Fuadi, K. (2013). *Tutorial matplotlib*. <http://twitter.com/sopier>.