

# Naskah Publikasi-Ellang Putro Priambodo-18.01.55.0039- 20062022 *by Tete Hayati*

---

**Submission date:** 20-Jun-2022 07:51AM (UTC+0700)

**Submission ID:** 1859695095

**File name:** Naskah\_Publikasi\_Ellang\_Putro\_Priambodo\_18.01.55.0039.pdf (547.8K)

**Word count:** 3201

**Character count:** 19819

## **Perbandingan Clustering Algoritma K-Means Dan Hierarchical Pada Stok Barang PT Multi Lestari**

**Ellang Putro Priambodo<sup>1\*</sup>, Arief Jananto<sup>2</sup>**  
 Program Studi Sistem Informasi, Universitas Stikubank Semarang  
 Jl. Tri Lomba Juang No.1, Mugassari, Semarang, Indonesia  
 \*e-mail Corresponding Author: ellang65.kaell@gmail.com

### *Abstrak*

*Ketersediaan merupakan komponen yang sangat penting. Manajemen yang baik adalah yang dapat memproses dan mengatur ketersediaan stok barang. Persediaan barang dilakukan supaya pada saat dibutuhkan barang tersebut ada. Pada penelitian ini penulis melakukan uji klastering pengelompokan barang berdasarkan jumlah stok awal, barang masuk, barang keluar, stok akhir dan jenis barang. data uji berjumlah 659 record dengan klasifikasi barang tinggi, sedang, rendah dan kosong. Algoritma menggunakan metode pengelompokan dengan teknik clustering yang umum digunakan yaitu algoritma K-Means dan menggunakan algoritma AHC sebagai metode pembandingan.*

**Kata kunci:** Data Mining; stok barang; kmeans; ahc

*Availability is a very important component. Good management is the one who can process and manage stock availability. Inventory of goods is carried out when needed the goods are there. In this study, the authors conducted a test of grouping goods based on the number of initial stock, incoming goods, outgoing goods, final stock and types of goods. The test data returned 659 records with the classification of high, medium, low and empty goods. The algorithm uses a grouping method with the commonly used clustering technique, namely the K-Means algorithm and uses the AHC algorithm as a comparison method.*

### **1. Pendahuluan**

PT Multi Lestari adalah perusahaan yang bergerak dibidang manufaktur yang memproduksi barang/ produk dengan standar keamanan dan mutu tinggi sesuai dengan Standar Nasional Indonesia (SNI) yang beralamat di Jl. Raya Semarang Demak KM 7 No 349, Trimulyo, Sayung, Kec. Sayung, Kota Semarang, Jawa Tengah 50117. Ketersediaan barang Multi Lestari merupakan komponen yang sangat penting. Sehingga sebagai manajemen yang baik dalam proses mengatur ketersediaan stok barang sangat diperlukan, untuk menghindari penumpukan barang yang sama dan barang yang kurang diminati oleh pembeli [1]. Ketersediaan barang yang tidak terkelola dengan baik juga berdampak pada keagenan ketika barang habis pada saat permintaan konsumen tinggi, yang akan terjadi adalah permintaan barang harus ditunda sehingga berdampak langsung pada penjualan barang. barang di instansi [2].

Berdasarkan uraian di atas, dan dari beberapa sumber referensi yang penulis baca baik itu dari jurnal atau penelitian sebelumnya, ada beberapa metode teknik data mining yang biasa digunakan untuk proses pengolahan data, seperti prediksi, estimasi, deskripsi, klasifikasi, asosiasi, dan pengklasteran [1].

Jenis clustering yang akan diteliti adalah hierarchical clustering dan k-means clustering, selain itu penelitian ini juga bertujuan untuk menunjukkan seberapa perbedaan data atau hasil antara hierarchical dan k-means clustering [3].

### **2. Tinjauan Pustaka**

Nengsi Anggraini, Jasmir, Pareza Alam Jusia [1], melakukan penelitian dengan judul "Penerapan Metode K-Means Clustering untuk Menentukan Persediaan Stok Barang Pada Toko Pemsart Jambi". Masalah pada penelitian ini yaitu belum adanya sebuah sistem manajemen

sehingga terjadinya stockout dan overstock. Penelitian memakai 2 tools data mining yaitu RapidMiner dan Weka.

Penelitian kedua dilakukan oleh Desy Exasanti [4]. dengan judul "Analisa Hasil Pengelompokan Wilayah Kejadian Non-Kebakaran Menggunakan Agglomerative Hierarchical Clustering Pada Dinas Pemadam Kebakaran Kota Semarang". Penelitian ini membahas tentang data-data yang ada secara khusus di Dinas Pemadam Kebakaran Kota Semarang yang dianalisis menggunakan data mining. Dalam penelitian ini penulis menggunakan algoritma AHC dengan tiga metode (Single Linkage, Average Linkage dan Complete Linkage ) menggunakan tool data mining WEKA. Dengan model analisa yang disusun menggunakan metode KDD (Knowledge Discovery in Databases).

Penelitian ketiga dilakukan oleh Rahmatika Diana Firdaus, Tri Ginanjar Laksana, Rima Dias Ramadani [5], mahasiswa Institut Teknologi Telkom Purwokerto yang berjudul "Pengelompokan Data Persediaan Obat Menggunakan Perbandingan Metode K-Means Dengan Hierarchical Clustering Single Linkage". masalah yang diteliti penulis yaitu sistem penanganan pengelompokan obat di Puskesmas II Ajibarang berdasarkan jenis obat masih dilakukan secara manual dan membutuhkan waktu.

Pada penelitian sebelumnya yang dari segi metode dan algoritma yang dilakukan oleh Nengsi Anggraini, Jasmir, Pareza Alam Jusia [1], menggunakan metode algoritma K-Means Clustering dan pada penelitian yang dilakukan oleh Penelitian kedua dilakukan oleh Desy Exasanti [4], Menggunakan Hierarchical Clustering. Serta pada penelitian berikutnya membandingkan antar dua algoritma, yaitu pada penelitian Tri Ginanjar Laksana, Rima Dias Ramadani [5], membandingkan algoritma K-Means clustering dan Hierarchical clustering.

Dari ketiga penelitian tersebut dapat ditarik kesimpulan bahwa data mining dapat mengklasterisasi atau mengelompokkan data. Banyak algoritma yang dapat digunakan untuk proses klusterisasi seperti K-Mean, K-Medoids dan Hierarchical juga tool yang digunakan untuk memprosesnya seperti RapidMiner, Rstudio dan WEKA.

Dalam penelitian ini, penulis menggunakan dua algoritma yaitu K-Means clustering dan Hierarchical clustering yang kemudian dibandingkan hasil proses dari kedua algoritma. Analisa dilakukan dengan mengelompokkan tiap data menjadi beberapa cluster kemudian hasil analisa digambarkan dalam bentuk tabel yang mana dalam prosesnya menggunakan tools Rstudio.

## 2. Metodologi

### 3.1 Algoritma K-Means

Algoritma K-Means adalah algoritma clustering iteratif yang mengelompokkan data menjadi sejumlah K cluster yang telah ditentukan. Algoritma K-Means mudah diterapkan dan dijalankan, relatif cepat, mudah beradaptasi dan umum digunakan.[8]

### 3.2 Algoritma Hierarchical

Pengelompokan hirarki (Hierarchical Clustering) adalah metode dengan cara melakukan pengelompokan objek-objek data kedalam sebuah kelompok hirarki [9]. Metode-metode Hierarchical clustering di kategorikan ke dalam agglomerative (bawah-atas) dan divisive (atas-bawah).

3.3 Tahapan Penelitian.



Gambar 1 Tahapan Penelitian

3.4 Metode Penelitian CRISP-DM

CRISP-DM (Cross-Industry Standard Process for Data Mining) adalah metodologi data mining yang dikembangkan oleh konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam data mining [6]. Deskripsi metode tidak bersifat teoritis/normatif, namun menyajikan prosedur atau tahapan-tahapan ril yang dilalui dalam upaya pencapaian tujuan akhir yang diharapkan dalam riset/penelitian [6]. Tahapan-tahapan metodologi CRISP-DM tersebut dijelaskan sebagai berikut :

1. Business Understanding

Pada penelitian datamining ini bertujuan membandingkan hasil cluster algoritma K-Means clustering dan Hierarchical clustering pada data stok barang.

2. Data Understanding

Data yang dipakai merupakan data stok barang pada PT Multi Lestari periode bulan Februari 2021 sampai bulan Juni 2021. potongan dataset asli PT Multi Lestari dapat dilihat pada Gambar 1.

1	2	3	4	5	6	7	TOTAL						STOK AKHIR			
							NAMA BARANG	SATUAN	STOK AKHIR	IN				SELISIH		
										STOCK AWAL	STOCK MASUK	END PROCESS		STOCK KELUAR	NG	WAJIB KEMBALI
8	Flexible Hose 1.5 M	PCS	-	60.250	93365	0	33115	0	33.115				60.250			
9	Hose Clamp 3/4	PCS	142.000	70.467	0	5631	77164	272	76.892				70.467			
10	HOSE CLAMP 7/8	PCS	55.652	55.652	0	0	0	0	0				55.652			
11	Karet Huk	PCS	100.000	90.930	0	0	69060	0	69.060				90.930			
12	Selang Hitam / Selang Gascamp Hitam 1.8 17 MM	PCS	6.250	2.050	28750	0	32990	0	22.100				2.050			
13	Selang Orange Gascamp High Pressure 1.8mm	PCS	-	0	0	0	0	0	14				0			
14	TALI PACKING	ROL	85	26	25	0	84	0	20				26			
15	Karet Huk Hitam/Hitam Plastik	PCS	70680	70.680	0	0	0	0	0				70.680			
16	HOSE CLAMP MODENA 3/4	PCS	98	3.496	4000	0	602	0	602				3.496			
17	LOGO															
18	LOGO ALLISON 32 PLAT	PCS	180	180	0	0	0	0	0				180			
19	LOGO ALLISON 32	PCS	-	0	0	0	0	0	0				0			
20	LOGO ALLISON 27 (stiker)	PCS	924	924	0	0	0	0	0				924			
21	LOGO ALLISON 27	PCS	-	0	0	0	0	0	0				0			
22	LOGO ALLISON 23 PLAT	PCS	63	63	0	0	0	0	0				63			
23	LOGO ALLISON 23	PCS	7	7	0	0	0	0	0				7			
24	LOGO ALLISON 20	PCS	928	896	0	0	120	0	120				896			
25	LOGO ALLISON 20 (stiker)	PCS	252	252	0	0	0	0	0				252			

Gambar 2 Potongan Dataset Asli PT Multi Lestari

3. Data Preparation

Pada tahap ini data akan di olah kembali karena pada data asli tidak semua data dan atribut akan di pakai pada penelitian. Berikut beberapa tahapan pada fase persiapan data :

a. Pemilihan Atribut

b. Pada tahap ini dilakukan seleksi pada atribut mana yang akan digunakan pada penelitian. Atribut hasil seleksi dapat dilihat pada Tabel 1.

Tabel 1 Atribut Hasil Seleksi

No	Atribut
1.	Nama_barang
2.	Stok_awal
3.	Stok_masuk
4.	Stok_keluar
5.	Stok_akhir

a. Pembersihan Data

Pada tahapan ini yaitu proses pembersihan data yaitu dengan menghapus data yang tidak lengkap, kosong, ataupun hilang dan menghapus data barang lama yang sudah di pakai.

b. Transformasi Data

Pada tahap ini akan di lakukan tranformasi data pada data Excel untuk mempermudah proses clustering.

1) Atribut Stok\_Awal, Barang\_Masuk, Barang\_Keluar, Stok\_Akhir. Hasil tranformasi atribut seleksi dapat dilihat pada Tabel 2.

Tabel 2 Tranformasi Atribut Hasil seleksi

atribut	Jumlah transaksi	klasifikasi	Konversi data
Stok_awal,	10.000 lebih	Tinggi	4
Barang_masuk,	5000-10.000	Sedang	3
Barang_keluar,	1-5000	Rendah	2
Stok_akhir	0	kosong	1

2) Jenis barang

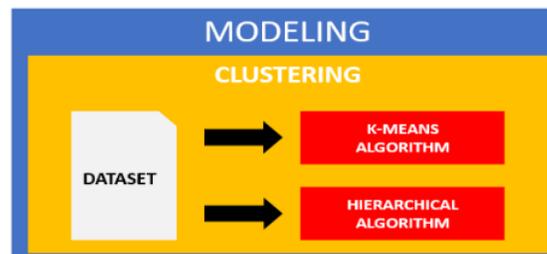
Merupakan atribut baru yang di buat untuk mengelompokan data pada tiap jenis barang. Dengan klasifikasi menjadi 7 kelompok yang dapat dilihat pada Tabel 3.

Tabel 3 Konversi Data Jenis\_barang

Jenis barang	Konversi data
packing	1
logo	2
blister	3
etiket	4
duplex	5
masterbox	6
komponen	7

4. Fase Pemodelan (Modeling)

Pada tahap pemodelan ini akan menggunakan dua algoritma data mining yaitu algoritma K-Means Clustering dan Hierarchical Clustering dengan tool datamining Rstudio yang kemudian akan didapatkan hasil kluster pada tiap stok barang. Fase pemodelan dapat dilihat pada Gambar 2.



Gambar 3 Fase Modeling penelitian

5. Evaluation  
Pada tahap ini melakukan evaluasi terhadap hasil dari proses data mining yang dihasilkan pada tahap pemodelan sebelumnya. Evaluasi apakah metode clustering yang menggunakan algoritma K-means clustering dan hierarchical clustering mendapat hasil yang sesuai dengan tujuan awal penelitian.
6. Deployment  
Tahap terakhir hasil proses data mining akan di implementasikan menjadi sebuah kesimpulan dalam bentuk laporan yang dibuat sedemikian rupa sehingga mudah dimengerti.

#### 4. Hasil dan Pembahasan

##### 4.1 Install package Rstudio

Beberapa packages yang diperlukan pada implementasi algoritma K-Means dan Hirarchical clustering agar dapat berjalan pada Rstudio yang perlu diinstall sebagai berikut :

```
Install.package("tidyverse")
library(tidyverse)
Install.package("cluster")
library(cluster)
Install.package("factoextra")
library(factoextra)
```

##### 4.2 Data Preparation

Tahap pertama adalah menyiapkan data yang akan diolah di Rstudio. Berikut sourcecode pada proses data preparation :

```
data_februari <- read_excel("februari_fx.xlsx")
```

Sourcecode yang berfungsi untuk membaca dan mengimport file "februari\_fx" yang berupa file excel yaitu data stok barang bulan februari kemudian ditempatkan pada variabel "data\_februari".

```
View(data_februari)
```

Sourcecode yang berfungsi melihat variabel data\_februari yang telah dibuat sebelumnya. Data bulan februari berjumlah 87 record dengan 6 kolom. Hasil import file dapat dilihat pada Gambar 3.

	Nama_Barang	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir	Jenis_barang
1	Hose Clamp 3/4	4	1	4	4	1
2	Karet Hek	4	1	4	4	1
3	Selang Hitam	3	4	4	2	1
4	TALI PACKING	2	2	2	2	1
5	HOSE CLAMP MODENA 3/4	2	2	2	2	1
6	LOGO ALLISON 20	2	1	2	2	2
7	LOGO LUXURY 27	2	1	2	2	2
8	LOGO LUXURY 20	2	1	2	2	2
9	LOGO KAZUKI 20	2	1	2	2	2
10	LOGO ZAMAN 20	2	2	2	2	2
11	LOGO QATAR 20	2	1	2	2	2
12	LOGO QUARDON 20	2	1	2	2	2
13	LOGO GASCOMP 20	2	4	3	4	2
14	LOGO GASCOMP 23	2	1	2	2	2
15	LOGO SUPERLOCK 20	4	1	2	4	2

Gambar 4Potongan Dataset Data Stok Barang Bulan Februari

```
summary(data_februari)
```

Sourcecode yang berfungsi melihat rincian data pada variabel data\_februari. Tampilan rincian data dapat dilihat pada Gambar 4.

```

> summary(data_februari)
  Nama_barang      Stok_awal      Stok_masuk      Stok_keluar
Length:87         Min.   :2.000   Min.   :1.000   Min.   :1.000
Class :character  1st Qu.:2.000   1st Qu.:1.000   1st Qu.:2.000
Mode  :character  Median :2.000   Median :1.000   Median :2.000
                               Mean  :2.862   Mean  :1.494   Mean  :2.598
                               3rd Qu.:4.000  3rd Qu.:1.000  3rd Qu.:4.000
                               Max.  :4.000   Max.  :4.000   Max.  :4.000

  Stok_akhir      Jenis_barang
Min.   :2.000     Min.   :1.000
1st Qu.:2.000     1st Qu.:2.000
Median :2.000     Median :5.000
Mean  :2.851     Mean  :4.517
3rd Qu.:4.000     3rd Qu.:7.000
Max.  :4.000     Max.  :7.000

```

Gambar 5 Rincian Data dari data\_februari.

```
februari <- na.omit(data_februari)
```

Soucecode yang berfungsi untuk menghapus jika terdapat nilai na's atau data yang kosong maupun hilang.

```
februari_data.numerik <- februari[2:5]
```

Soucecode yang berfungsi untuk mengambil data numerik yang berada pada kolom kedua sampai kolom kelima.

```
februari_datafix <- scale(februari_data.numerik)
```

Soucecode yang berfungsi untuk transformasi skala data\_februari. Transformasi data numerik dilakukan dengan standarisasi data. Jika rentang nilai antar variabel memiliki perbedaan skala yang cukup besar yang dapat menyebabkan bias dalam analisis cluster. Hasil tranformasi skala dapat dilihat pada Gambar 5.

	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir
1	1.2238827	-0.4838248	1.5442117	1.2142905
2	1.2238827	-0.4838248	1.5442117	1.2142905
3	0.1483494	2.4528794	1.5442117	-0.8985750
4	-0.9271838	0.4950766	-0.6581886	-0.8985750
5	-0.9271838	0.4950766	-0.6581886	-0.8985750
6	-0.9271838	-0.4838248	-0.6581886	-0.8985750
7	-0.9271838	-0.4838248	-0.6581886	-0.8985750
8	-0.9271838	-0.4838248	-0.6581886	-0.8985750
9	-0.9271838	-0.4838248	-0.6581886	-0.8985750
10	-0.9271838	0.4950766	-0.6581886	-0.8985750
11	-0.9271838	-0.4838248	-0.6581886	-0.8985750
12	-0.9271838	-0.4838248	-0.6581886	-0.8985750
13	-0.9271838	2.4528794	0.4430116	1.2142905
14	-0.9271838	-0.4838248	-0.6581886	-0.8985750
15	1.2238827	-0.4838248	-0.6581886	1.2142905

Gambar 6 Potongan Hasil Transformasi Skala.

### 4.3 Eksekusi Algoritma K-means

Dalam proses clustering jumlah cluster yang dibentuk yaitu tiga cluster. Berikut soucecode yang digunakan dalam proses perhitungan algoritma K-Mean pada Rstudio.

```
februari_final <- kmeans(februari_datafix, centers = 3)
```

Soucecode yang berfungsi untuk memproses clusterisasi data\_februari dengan jumlah tiga cluster. Hasil clustering kemudian akan ditempatkan pada variabel "februari\_final".

```
print(februari_final)
```

Soucecode yang berfungsi untuk menampilkan hasil proses clustering pada proses sebelumnya. Hasil perhitungan dapat dilihat pada Gambar 6 dan cluster vector dan cluster sum of squares dapat dilihat pada Gambar 7.

```

K-means clustering with 3 clusters of sizes 16, 46, 25

Cluster means:
  Stok_awal Stok_masuk Stok_keluar Stok_akhir
1  1.0222202 -0.4838248 -0.4517136  0.9501823
2 -0.8336592 -0.2710202 -0.6581886 -0.8756090
3  0.8797120  0.8083250  1.5001637  1.0030040

```

Gambar 7 Tampilan Cluster Means.

Pada gambar tersebut dapat diketahui jumlah cluster yang terbentuk berjumlah tiga cluster dengan anggota cluster pertama berjumlah 16, anggota cluster kedua berjumlah 46 dan anggota cluster ketiga berjumlah 25. Cluster mean dari setiap atribut dikelompokkan menjadi tiga cluster.

```

Clustering vector:
[1] 3 3 3 2 2 2 2 2 2 2 2 2 2 2 3 2 1 2 2 2 2 1 2 1 1 2 2 2 2 2 2 2 3 1 2 2 2
[35] 2 1 2 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3
[69] 3 3 3 3 2 2 2 3 2 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Within cluster sum of squares by cluster:
[1] 9.123596 17.157631 71.905629
(between_SS / total_SS = 71.5 %)

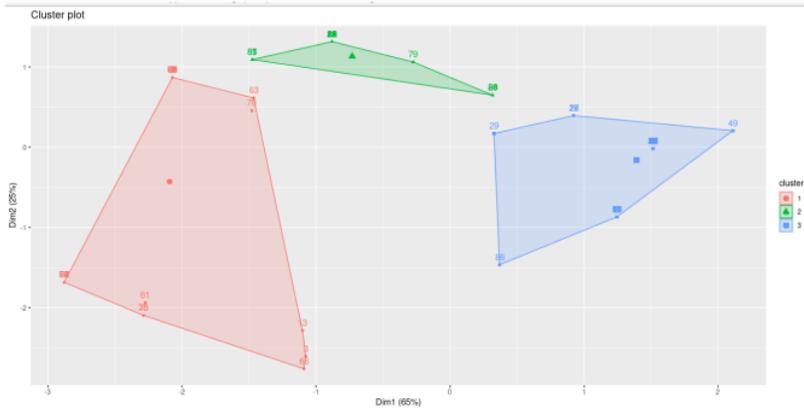
```

Gambar 8 Tampilan Clustering Vector dan Cluster Sum Of Squares.

Pada gambar tersebut dapat diketahui bahwa pengelompokan dengan 86 record pada tiap cluster dan ditampilkan pada sebuah clustering vector. Nilai sum of squares pada cluster satu 9.123596, cluster kedua bernilai 17.157631 dan cluster ketiga bernilai 71.905629. Dengan nilai perhitungan pembagian between sum of squares 71.5%.

```
fviz_cluster(februari_final,data=februari_datafix)
```

Soucecode yang berfungsi untuk visualisasi hasil analisa perhitungan k-mean dalam grafik. Hasil visualisasi dapat dilihat pada Gambar 8.



Gambar 9 Grafik Plot Visualisasi K-means.

```
februari_finalakhir=data.frame(februari,februari_final$cluster)
view(februari_finalakhir)
```

Soucecode yang berfungsi untuk menampilkan hasil clustering data dalam tabel. Hasil visualisasi K-Means dalam bentuk tabel dapat dilihat pada gambar 9.

	Nama_Barang	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir	Jenis_barang	februari_final.cluster
1	Hose Clamp 3/4	4	1	4	4	1	3
2	Karet Hek	4	1	4	4	1	3
3	Selang Hitam	3	4	4	2	1	3
4	TALI PACKING	2	2	2	2	1	2
5	HOSE CLAMP MODENA 3/4	2	2	2	2	1	2
6	LOGO ALLISON 20	2	1	2	2	2	2
7	LOGO LUXURY 27	2	1	2	2	2	2
8	LOGO LUXURY 20	2	1	2	2	2	2
9	LOGO KAZUKI 20	2	1	2	2	2	2
10	LOGO ZAMAN 20	2	2	2	2	2	2
11	LOGO QATAR 20	2	1	2	2	2	2
12	LOGO QUARDON 20	2	1	2	2	2	2
13	LOGO GASCOMP 20	2	4	3	4	2	3
14	LOGO GASCOMP 23	2	1	2	2	2	2
15	LOGO SUPERLOCK 20	4	1	2	4	2	1

Gambar 10 Visualisasi K-Means Clustering dalam Bentuk Tabel.

```
februari[2:5] %>%
mutate(Cluster=februari_final$cluster) %>%
group_by(Cluster) %>%
summarise_all("mean")
```

Soucecode yang berfungsi untuk menampilkan ringkasan rata-rata dari masing-masing cluster pada tiap atribut. Hasil dapat dilihat pada Gambar 10.

	Cluster	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir
	<int>	<dbl>	<dbl>	<dbl>	<dbl>
1	1	3.81	1	2.19	3.75
2	2	2.09	1.22	2	2.02
3	3	3.68	2.32	3.96	3.8

Gambar 11 Ringkasan Hasil Clustering.

```
write_xlsx(x = februari_finalakhir, path = "kmean_februari_final.xlsx", col_names = TRUE)
```

Soucecode yang berfungsi untuk mengexport hasil akhir K-Means clustering dalam bentuk file excel.

#### 4.4 Eksejusi Algoritma Hierarchical

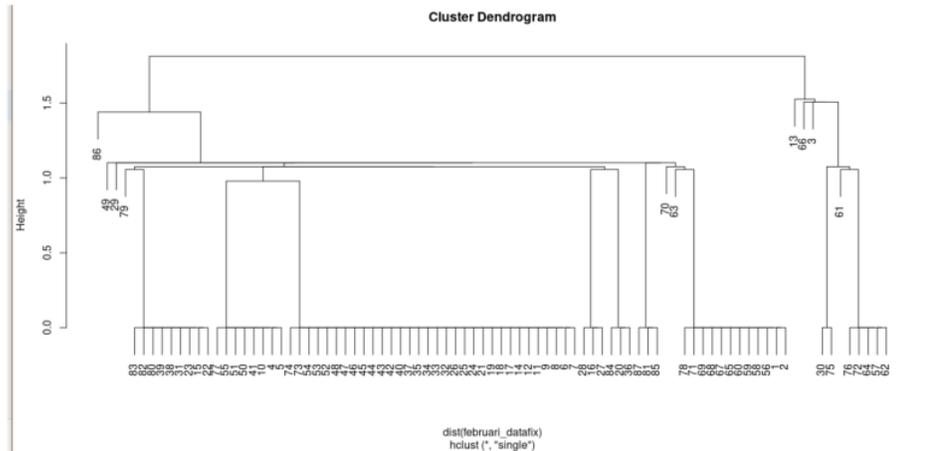
Dalam proses clustering jumlah cluster yang dibentuk yaitu tiga cluster. Berikut soucecode yang digunakan dalam proses perhitungan algoritma Hierarchical pada Rstudio.

```
hc_februari <- hclust(dist(februari_datafix), method="single")
```

Soucecode yang berfungsi untuk memproses clusterisasi dengan algoritma Hierarchical metode single link pada data\_februari. Hasil clustering kemudian akan ditempatkan pada variabel "hc\_februari".

```
plot(hc_februari)
```

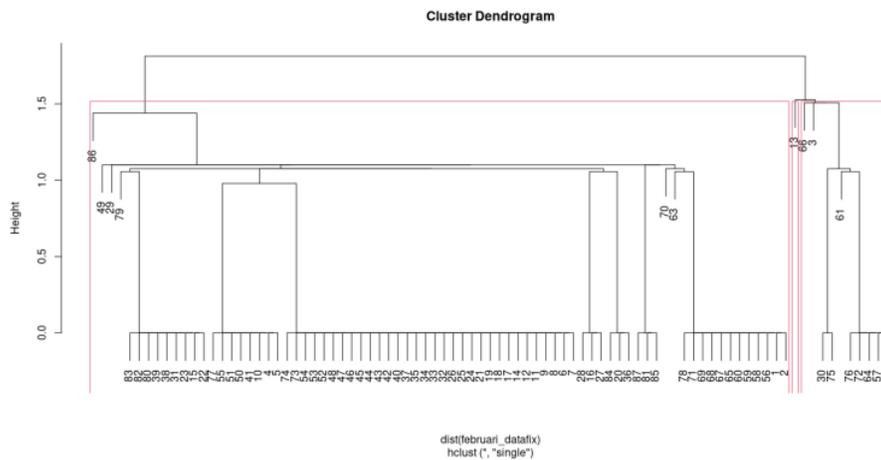
Soucecode yang berfungsi untuk menampilkan hasil grafik plot algoritma Hierarchical atau dendrogram berdasarkan atribut "No" pada data\_februari. Hasil dapat dilihat pada Gambar 12.



Gambar 12 Dendrogram Hirarchical Clustering.

```
rect.hclust(hc_februari,k=3)
groups<-cutree(hc_februari,k=3)
```

Soucecode yang berfungsi untuk membentuk 3 cluster. Hasil dapat dilihat pada Gambar 13.



Gambar 13 Dendrogram Hirarchical Clustering dengan Tiga Cluster.

```
februari_finalakhir=data.frame(februari, final)
view(februari_finalakhir)
```

Soucecode yang berfungsi untuk menampilkan hasil clustering data dalam tabel. Hasil visualisasi Hirarchical Clustering dalam bentuk tabel dapat dilihat pada gambar 14.

	Nama_Barang	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir	Jenis_barang	groups
1	Hose Clamp 3/4	4	1	4	4	1	1
2	Karet Hek	4	1	4	4	1	1
3	Selang Hitam	3	4	4	2	1	2
4	TALI PACKING	2	2	2	2	1	1
5	HOSE CLAMP MODENA 3/4	2	2	2	2	1	1
6	LOGO ALLISON 20	2	1	2	2	2	1
7	LOGO LUXURY 27	2	1	2	2	2	1
8	LOGO LUXURY 20	2	1	2	2	2	1
9	LOGO KAZUKI 20	2	1	2	2	2	1
10	LOGO ZAMAN 20	2	2	2	2	2	1
11	LOGO QATAR 20	2	1	2	2	2	1
12	LOGO QUARDON 20	2	1	2	2	2	1
13	LOGO GASCOMP 20	2	4	3	4	2	3
14	LOGO GASCOMP 23	2	1	2	2	2	1
15	LOGO SUPERLOCK 20	4	1	2	4	2	1

Showing 1 to 15 of 87 entries. 7 total columns

**Gambar 14** Visualisasi Hirarchical Clustering dalam Bentuk Tabel.

```
write_xlsx(x = februari_finalakhir, path = "kmean_februari_final.xlsx", col_names = TRUE)
```

Soucecode yang berfungsi untuk mengexport hasil akhir Hirarchical clustering dalam bentuk file excel.

#### 4.5 Pembahasan dan Hasil Analisa

Setelah melakukan perhitungan Algoritma K-Means dan Hirarchical pada bulan Februari sampai bulan Juni dengan Rstudio diperoleh hasil yang dapat dilihat pada Tabel 4.

*Tabel 4 Hasil Clustering Berdasarkan Jumlah Anggota Cluster.*

Bulan	Cluster	K-mean	Hirarchical
Februari	1	16	76
	2	46	10
	3	25	1
Maret	1	27	86
	2	8	18
	3	72	3
April	1	27	83
	2	29	20
	3	48	1
Mei	1	24	22
	2	53	83
	3	30	2
Juni	1	25	14
	2	92	133
	3	32	2

Dari Tabel 4 diketahui bahwa hasil clustering algoritma K-Means dan Hirarchical mempunyai hasil yang berbeda. Contoh pada bulan Februari algoritma K-Mean pengelompokan cluster ke-1 dengan jumlah anggota 12, cluster ke-2 dengan jumlah anggota 46 dan cluster ke-3 dengan jumlah anggota 25. Sedangkan pada algoritma Hirarchical mengelompokan cluster ke-1 dengan jumlah anggota 76, cluster ke-2 dengan jumlah anggota 10 dan cluster ke-3 dengan jumlah anggota 1.

Dari hasil clustering algoritma K-Means dan Hirarchical diketahui bahwa pembagian anggota cluster algoritma K-mean lebih seimbang. Sedangkan pembagian anggota cluster algoritma Hirarchical terjadi ketidak seimbangan yaitu jumlah anggota cluster ke-1 atau cluster ke-2 akan lebih banyak dari cluster ke-3.

Dari hasil clustering pada Tabel 4 dijabarkan kembali menjadi hasil rekap jumlah anggota cluster berdasarkan jenis barang. Hasil dapat dilihat pada Gambar 15.

Bulan	Cluster	K-Means							Hirarchical								
		packing	logo	blister	etiket	duplex	masterbox	komponen	jumlah	packing	logo	blister	etiket	duplex	masterbox	komponen	jumlah
Februari	1	0	2	2	2	2	0	8	16	11	15	3	11	4	8	24	76
	2	9	13	1	9	2	8	4	46	1	0	0	1	0	0	8	10
	3	3	1	0	1	0	0	20	25	0	1	0	0	0	0	0	1
Maret	1	3	1	1	1	0	0	21	27	11	15	2	11	10	9	28	86
	2	0	1	0	3	0	0	4	8	2	0	1	1	0	0	14	18
	3	10	15	2	8	10	9	18	72	0	1	0	0	0	0	2	3
April	1	0	2	1	1	3	0	20	27	10	11	2	9	12	10	29	83
	2	2	3	1	1	2	0	20	29	1	0	1	1	0	0	17	20
	3	9	7	1	8	7	10	6	48	0	1	0	0	0	0	0	1
Mei	1	2	1	1	0	0	0	20	24	1	0	1	0	0	0	20	22
	2	9	13	0	10	8	10	3	53	10	15	1	12	11	10	24	83
	3	1	2	1	2	3	0	21	30	1	1	0	0	0	0	0	2
Juni	1	0	2	2	3	5	0	13	25	1	0	0	1	0	0	12	14
	2	7	16	3	28	21	13	4	92	9	19	6	30	27	13	29	133
	3	3	1	1	1	1	0	25	32	0	0	0	1	0	0	1	2

Gambar 15 Hasil Rekap Jumlah Anggota Clustering Berdasarkan Jenis Barang.

Berdasarkan hasil pada Gambar 14 diketahui bahwa jenis barang dengan jumlah tertinggi pada algoritma K-Mean terjadi pada bulan Juni di cluster ke-2 dengan jenis barang etiket dan jumlah 28 item. Sedangkan pada algoritma Hirarchical terjadi pada bulan Juni di cluster ke-2 dengan jenis barang etiket dan jumlah 30 item

Dari hasil clustering pada Tabel 4 dijabarkan kembali menjadi hasil rekap jumlah anggota cluster berdasarkan kategori transaksi dengan contoh bulan Juni. Hasil dapat dilihat pada Gambar 4.16.

Bulan	Cluster	Kategori	K-Means				Hirarchical			
			Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir	Stok_awal	Stok_masuk	Stok_keluar	Stok_akhir
Juni	cluster 1	tinggi	25	2	0	23	10	14	13	11
		sedang	0	0	11	2	2	0	1	2
		rendah	0	0	12	0	2	0	0	1
	cluster 2	kosong	0	23	2	0	0	0	0	0
		tinggi	0	0	0	0	38	3	17	36
		sedang	12	6	2	10	12	6	13	13
	cluster 3	rendah	80	14	88	82	83	14	101	84
		kosong	0	72	2	0	0	110	2	0
		tinggi	25	17	30	26	2	2	0	2
		sedang	2	0	1	3	0	0	0	0
		rendah	5	0	1	3	0	0	0	0
		kosong	0	15	0	0	0	0	2	0

Gambar 16 Hasil Rekap Jumlah Anggota Clustering Berdasarkan Kategori Transaksi bulan Juni.

Dapat diketahui pada algoritma K-Mean Stok\_awal tertinggi terjadi pada bulan juni cluster ke-2 dengan kategori rendah dengan jumlah 80. Sedangkan pada algoritma Hirarchical Stok\_awal tertinggi terjadi pada bulan Juni cluster ke-2 dengan kategori rendah dengan jumlah 83.

Dapat diketahui pada algoritma K-Mean Stok\_akhir tertinggi terjadi pada bulan juni cluster ke-2 dengan kategori rendah dengan jumlah 82. Sedangkan pada algoritma Hirarchical Stok\_awal tertinggi terjadi pada bulan Juni cluster ke-2 dengan kategori rendah dengan jumlah 84.

Dari hasil rekap jumlah anggota clustering berdasarkan kategori transaksi algoritma K-means dan Hirarchical mempunyai kemiripan pada transaksi tertinggi.

## 5. Simpulan

Dari hasil penelitian pada data Stok barang PT Multi Lestari menggunakan Algoritma K-Mean dan Hierarchical dapat disimpulkan sebagai berikut :

1. Algoritma K-Means dan Hierarchical telah berhasil diterapkan pada data stok barang PT Multi Lestari.
2. Berdasarkan jumlah anggota cluster algoritma K-mean mempunyai pembagian anggota yang lebih seimbang. Sedangkan pembagian anggota cluster algoritma Hierarchical terjadi ketidak seimbangan yaitu jumlah anggota cluster ke-1 atau cluster ke-2 akan lebih banyak dari cluster ke-3. Contoh pada bulan Februari Algoritma K-mean pada cluster ke-1 berjumlah 16, cluster ke-2 berjumlah 46 dan cluster ke-3 berjumlah 25, sedangkan algoritma hierarchical pada cluster ke-1 berjumlah 76, cluster ke-2 berjumlah 10 dan cluster ke-3 berjumlah 1.
3. Pembagian anggota cluster algoritma K-Means dan Hierarchical mempunyai hasil yang jauh berbeda, tetapi pada hasil rekap jumlah anggota clustering berdasarkan jenis barang dan hasil rekap jumlah anggota clustering berdasarkan kategori transaksi mempunyai kemiripan pada hasil yang tertinggi. Contoh pada rekap jenis barang hasil tertinggi algoritma K-Mean terjadi pada bulan Juni di cluster ke-2 dengan jenis barang etiket dan jumlah 28 item. Sedangkan pada algoritma Hierarchical terjadi pada bulan Juni di cluster ke-2 dengan jenis barang etiket dan jumlah 30 item.

**Daftar Referensi**

- [1] N. Anggraini, J. Jasmir, and P. A. Jusia, "Penerapan Metode K-Means Clustering untuk Menentukan Persediaan Stok Barang Pada Toko Pensmart Jambi," *J. Ilm. Mhs.*, pp. 63–77, 2019
- [2] S. Kasus and K. Ben, "Proceeding SENDIU 2021 PERBANDINGAN CLUSTERING OPTIMALISASI STOK BARANG MENGGUNAKAN ALGORITMA K – MEANS DAN ALGORITMA K – MEDOIDS," pp. 978–979, 2021.
- [3] H. Setiawan, K. Karamitta, K. Hosea, and L. Hakim, "Analisa Perbandingan Metode Hierarchical dan K-means dalam Clustering Data Terhadap Penjualan Jajansamavivi Comparative Analysis of Hierarchical and K-means Methods in Clustering Data on Sales of Jajansamavivi," pp. 91–99, 2020.
- [4] D. Exasanti and A. Jananto, "Analisa Hasil Pengelompokan Wilayah Kejadian Non-Kebakaran Menggunakan Agglomerative Hierarchical Clustering pada Dinas Pemadam Kebakaran Kota Semarang," vol. 15, no. 2, pp. 63–75.
- [5] R. D. Firdaus, T. G. Laksana, and R. D. Ramadhani, "Pengelompokan Data Persediaan Obat Menggunakan Perbandingan Metode K-Means Dengan Hierarchical Clustering Single Linkage," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 2, no. 1, pp. 33–48, 2019, doi: 10.20895/inista.v2i1.87.
- [6] I. Budiman, T. Prahasto, and Y. Christyono, "Data Clustering Menggunakan Metodologi CRISP-DM Untuk Pengenalan Pola Proporsi Pelaksanaan Tridharma," *J. Sist. Inf. Bisnis*, vol. 1, no. 3, pp. 15–16, 2014, doi: 10.21456/vol1iss3pp129-134.
- [7] R. Setiawan, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru ( Studi Kasus : Politeknik Lp3i Jakarta )," *J. Lentera Ict*, vol. 3, no. 1, pp. 76–92, 2016.
- [8] U. N. Putra *et al.*, "Penerapan Data Mining K-Means Clustering Untuk," pp. 241–249, 2021.
- [9] I. H. Witten, E. Frank, dan M. A. Hall, *Data mining: practical machine learning tools and techniques*, 3rd ed. Burlington, MA: Morgan Kaufmann, 2011.
- [10] R. D. Firdaus, T. G. Laksana, and R. D. Ramadhani, "Pengelompokan Data Persediaan Obat Menggunakan Perbandingan Metode K-Means Dengan Hierarchical Clustering Single Linkage," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 2, no. 1, pp. 33–48, 2019, doi: 10.20895/inista.v2i1.87.
- [11] T. H. Sardar and Z. Ansari, "An analysis of MapReduce efficiency in document clustering using parallel K-means algorithm," *Futur. Comput. Informatics J.*, pp. 1–10, 2018.
- [12] Sulastri H, Gufroni AI. Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia. *Jurnal Teknologi dan Sistem Informasi*. 2017 Sep 26;3(2):299-305.

# Naskah Publikasi-Ellang Putro Priambodo-18.01.55.0039-20062022

## ORIGINALITY REPORT

7%

SIMILARITY INDEX

4%

INTERNET SOURCES

0%

PUBLICATIONS

4%

STUDENT PAPERS

## PRIMARY SOURCES

1

Submitted to Sekolah Tinggi Sandi Negara

Student Paper

3%

2

doku.pub

Internet Source

2%

3

[ejurnal.teknokrat.ac.id](http://ejurnal.teknokrat.ac.id)

Internet Source

2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On