

# KLASTERISASI UJI LABORATORIUM SEBAGAI PENUNJANG DIAGNOSA PENYAKIT DEMAM BERDARAH MENGGUNAKAN K-MEANS

Sintia Budianti Ningsih<sup>1)</sup>, Agung Nilogiri<sup>2)</sup>  
Program Studi Teknik Informatika, Fakultas Teknik  
Universitas Muhammadiyah Jember  
Jl. Karimata No. 49 Jember Kode Pos 68161  
Email : [sintia2909.sa@gmail.com](mailto:sintia2909.sa@gmail.com)

## ABSTRAK

Demam berdarah merupakan penyakit yang disebabkan oleh tipe infeksi virus *dengue* yang ditularkan melalui gigitan nyamuk *aedes aegypti*. Virus *dengue* merupakan anggota *family flaviridae*. Rumah Sakit Citra Husada menangani berbagai macam penyakit termasuk penyakit Demam Berdarah. Namun penanganan pasien dilakukan secara individual dan belum ada pengelompokkan berdasarkan parameter gejala. *Clustering* adalah metode yang digunakan dalam data mining yang cara kerjanya mencari dan mengelompokkan data satu dengan data lainnya yang telah diperoleh. *K-Means* bertujuan untuk membuat *cluster* objek berdasarkan atribut menjadi *k* partisi. Penelitian ini mencari nilai *k* terbaik menggunakan metode *elbow*. Untuk hasil pengelompokkan menggunakan *K-Means* didapatkan jumlah data pasien di masing-masing *cluster* yaitu *cluster* 1 terdapat 47 data pasien, *cluster* 2 terdapat 30 data pasien, dan *cluster* 3 terdapat 8 data pasien.

Kata Kunci : Demam Berdarah, *K-Means*, *Elbow*

## I. PENDAHULUAN

Demam berdarah merupakan penyakit yang disebabkan oleh tipe infeksi virus *dengue* yang ditularkan melalui gigitan nyamuk *Aedes aegypti*. Virus *dengue* merupakan anggota *famili flaviridae*. Beberapa tipe virus *dengue* menunjukkan banyak persamaan karakteristik dengan *flavivirus* yang lainnya (Ginanjari, 2008). Karena banyaknya persamaan karakteristik tersebut, maka perlu adanya pengelompokkan/peng-*cluster*-an. Dimana pengelompokkan tersebut untuk melihat perbedaan karakteristik pasien yang nantinya ada dalam tabel kolom hasil pengelompokkan/peng-*cluster*-an.

Rumah Sakit Citra Husada Jember merupakan rumah sakit swasta yang didirikan pada tanggal 16 April 2009. Rumah Sakit Citra Husada diharapkan mampu menjawab keinginan masyarakat, khususnya masyarakat jember terhadap pelayanan kesehatan yang bermutu. Baik dari segi pelayanan kesehatan, sumber daya kesehatan yang profesional, sarana dan prasarana yang memenuhi standart kenyamanan, tenang dan keamanan terjamin. Rumah Sakit Citra Husada juga menangani berbagai macam

penyakit termasuk penyakit Demam Berdarah. Namun penanganan pasien dilakukan secara individual dan belum ada pengelompokkan berdasarkan parameter gejala.

*Clustering* adalah metode yang digunakan dalam data mining yang cara kerjanya mencari dan mengelompokkan data satu dengan data lainnya yang telah diperoleh. Ciri khas dari teknik data mining ini adalah mempunyai sifat tanpa arahan (*unsupervised*), yang dimaksud adalah teknik ini diterapkan tanpa perlu data *training* dan tanpa ada *teacher* serta tidak memerlukan target *output* (J.O. Ong, 2013). Metode *clustering* yang mempunyai sifat efisien dan cepat yang dapat digunakan salah satunya adalah metode *K-Means*, metode ini bertujuan untuk membuat *cluster* objek berdasarkan atribut menjadi *k* partisi. Cara kerja metode ini adalah mula-mula ditentukan *cluster* yang akan dibentuk, pada elemen pertama dalam tiap *cluster* dapat dipilih untuk dijadikan sebagai titik tengah (*centroid*), selanjutnya akan dilakukan pengulangan langkah-langkah hingga tidak ada objek yang dapat dipindahkan lagi (Y. Ardihila, H. Tjandra dan I. Arieshanti, 2014).

Identifikasi jumlah cluster  $k$  merupakan yang paling penting dan utama pada proses clustering dengan menggunakan algoritma  $K$ -Means dimana hasil cluster akan bergantung pada jumlah *cluster* awal. Sehingga jika jumlah cluster yang ditentukan tidak baik maka hasil cluster juga tidak akan sesuai dengan yang diharapkan yaitu tidak akan menghasilkan informasi yang dibutuhkan oleh pengguna. Maka perlu dilakukan penelitian dengan mencari nilai  $k$  terbaik menggunakan metode salah satunya adalah *elbow*. Metode ini akan melihat fungsi dari nilai *cluster* pada suatu data (Bholowalia, P. & Kumar, A., 2014). Metode *elbow* mudah diimplementasikan dengan cara melihat grafik dari nilai  $k$  yang akan diinputkan. Nilai fungsi  $k$  yang akan dibandingkan pada metode *elbow* adalah dengan melihat nilai SSE (*Sum of Square Error*) pada nilai *cluster* yang ditentukan. Jumlah  $k$  *cluster* terbaik akan dijadikan proses dari akhir perhitungan metode  $K$ -Means.

## II. Clustering

*Clustering* atau pengklasteran adalah suatu teknik *data mining* yang digunakan untuk menganalisis data untuk memecahkan permasalahan dalam pengelompokan data atau lebih tepatnya mempartisi dari *dataset* ke dalam subset. Pada teknik *clustering* targetnya adalah untuk kasus pendistribusian (objek, orang, peristiwa dan lainnya) ke dalam suatu kelompok, hingga derajat tingkat keterhubungan antar anggota *cluster* yang sama adalah kuat dan lemah antara anggota *cluster* yang berbeda (S. Agustina, D. Yudho, H. Santoso, dkk, 2012).

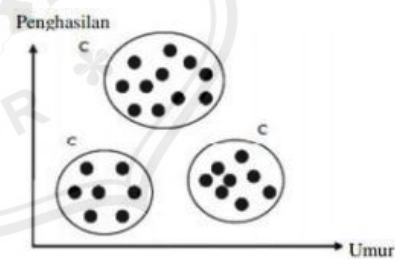
Teknik *cluster* mempunyai dua metode dalam pengelompokannya yaitu *hierarchical clustering* dan *non-hierarchical clustering*. *Hierarchical clustering* merupakan suatu metode pengelompokan data yang cara kerjanya dengan mengelompokkan dua data atau lebih yang mempunyai kesamaan atau kemiripan, kemudian proses dilanjutkan ke objek lain yang memiliki kedekatan dua, proses ini terus berlangsung hingga *cluster* membentuk

semacam *tree* dimana ada hirarki atau tingkatan yang jelas antar objek dari yang paling mirip hingga yang paling tidak mirip. Namun secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Sedangkan *non-hierarchical clustering* pada teknik ini dimulai dengan menentukan jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, empat *cluster*, atau lebih), setelah jumlah *cluster* yang diinginkan maka proses *cluster* dimulai tanpa mengikuti proses hirarki, metode ini juga sering disebut sebagai metode  $K$ -Means *clustering* (J. O. Ong, 2013).

Pada proses analisis *cluster* metode yang digunakan untuk membagi data menjadi subset data berdasarkan kesamaan atau kemiripan yang telah ditentukan sebelumnya. Jadi analisis *cluster* secara umum dapat dikatakan bahwa (S. Agustina, D. Yudho, H. Santoso, dkk, 2012):

1. Data yang terdapat dalam satu *cluster* memiliki tingkat kesamaan yang tinggi,
2. Dan yang terdapat dalam suatu *cluster* yang berbeda memiliki tingkat kesamaan yang rendah

Sebagai contoh dapat dilihat pada gambar 2.1 di bawah ini :



Gambar 1 Grafik *Clustering*

Pada gambar 1 dapat dilihat kita misalkan data tersebut merupakan data konsumen sederhana yang terdapat dua atribut didalamnya, yaitu umur dan penghasilan. Pada data yang berdasarkan dua atribut tersebut kemudian dibagi menjadi tiga *cluster* yaitu *cluster C1* yang terdiri dari konsumen usia muda dan berpenghasilan rendah, *cluster C2* terdiri dari konsumen usia muda dan tua berpenghasilan tinggi, dan

cluster C3 terdiri dari konsumen usia tua dan berpenghasilan relatif rendah.

### III. Metode K-Means

Algoritma *K-Means* merupakan salah satu algoritma dengan *partitional*, karena *K-Means* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya (Madhulata, T.S., 2012). Algoritma *K-Means* menggunakan proses secara berulang-ulang untuk mendapatkan basis data *cluster*. Dibutuhkan jumlah *cluster* awal yang diinginkan sebagai masukan dan menghasilkan jumlah *cluster* akhir sebagai *output*. Jika algoritma diperlukan untuk menghasilkan *cluster* K maka akan ada K awal dan K akhir. Metode *K-Means* akan memilih pola K sebagai titik awal *centroid* secara acak. Jumlah iterasi untuk mencapai *cluster centroid* awal. Akan dihitung dengan data/objek. Data yang memiliki jarak pendek atau terdekat dengan *centroid* akan membentuk sebuah *cluster* (Agrawal, A. & Gupta, H., 2013).

Prinsip kerja dari pengelompokan *hierarchical clustering* dilakukan secara bertahap. Dan disetiap iterasi dari pengelompokan *hierarchical clustering* hanya ada satu pemilihan penggabungan suatu item terhadap item lainnya (Y. Siyamto, 2017). Proses dasar algoritma *K-Means* dapat dilihat di bawah ini:

1. Tentukan jumlah *cluster* yang ingin dibentuk dan tetapkan pusat *cluster* k
2. Menggunakan jarak *euclidean* kemudian hitung setiap data ke pusat *cluster*

$$D_{L_2} = \sqrt{\sum_{j=1}^p (x_{2j} - x_{1j})^2}$$

3. Kelompokkan data ke dalam *cluster* dengan jarak yang paling pendek dengan persamaan

$$\text{Min} \sum_k^k = d_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}$$

4. Hitung pusat *cluster* yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p}$$

Dengan:

$x_{ij}$  = *cluster* ke-k

$p$  = banyaknya anggota *cluster* ke-k

Ulangi langkah dua sampai dengan empat sehingga sudah tidak ada lagi data yang berpindah ke *cluster* yang lain (S. Agustina, D. Yhudo, H. Santoso, dkk, 2012).

### IV. Metode Elbow

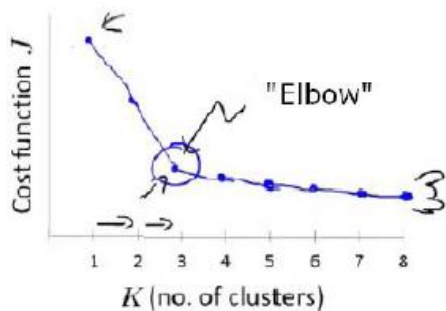
Metode *elbow* merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik (Madhulatha, T.S., 2012). Metode ini memebrikan ide/gagasan dengan cara memilih nilai *cluster* dan kemudian menambah nilai *cluster* tersebut untuk dijadikan model data dalam penentuan *cluster* terbaik. Dan selain itu persentase perhitungan yang dihasilkan menjadi perbandingan antara jumlah *cluster* yang ditambah (Kodinariya, Trupti M. & Makwana, Prashant R., 2013). Hasil persentase yang berbeda dari setiap nilai *cluster* dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika, nilai *cluster* kedua memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka nilai *cluster* tersebut yang terbaik (Bholowalia, Purnima & Kumar, Arvind, 2014).

Untuk mendapatkan perbandingannya adalah dengan menghitung SSE (*Sum of Square Error*) dari masing-masing nilai *cluster*. Karena semakin besar jumlah *cluster* K maka nilai SSE akan semakin kecil. Rumus SSE pada *K-Means* (Irwanto, et. Al, 2012):

$$SSE = \sum_{k=1}^K \sum_{x_i \in S_k} ||x_i - c_k||_2^2$$

Setelah dilihat akan ada beberapa nilai K yang mengalami penurunan paling besar dan selanjutnya hasil dari nilai K akan turun secara perlahan-lahan sampai hasil dari nilai K tersebut stabil. Misalnya nilai *cluster* K=2 ke K=3, kemudian dari K=3 ke K=4, terlihat penurunan drastis membentuk siku pada titik K=3 maka nilai *cluster* k yang ideal adalah

K=3 (Kodinariya, Trupti M. & Makwana, Prashant R., 2013).



Gambar 2.2 Grafik Metode *Elbow*  
 Algoritma Metode *Elbow* dalam menentukan nilai K pada *K-Means*:

1. Mulai
2. Inisialisasi awal nilai K
3. Naikkan nilai K
4. Hitung hasil *sum of square error* dari tiap nilai K
5. Melihat hasil *sum of square error* dari nilai K yang turun secara drastis
6. Tetapkan nilai K yang berbentuk siku
7. Selesai (Bholowalia, Purnima & Kumar, Arvind, 2014)

## V. Implementasi dan Pengujian Implementasi Metode *K-Means*

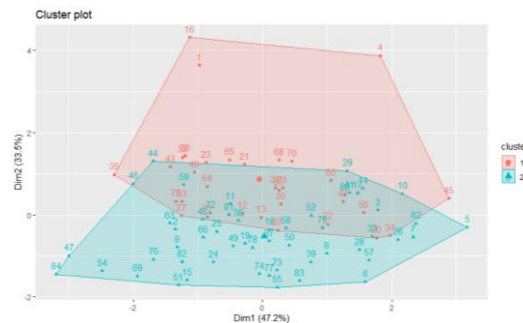
Dari 85 data pasien yang akan dikelompokkan menjadi beberapa *cluster* menggunakan algoritma *k-means* pada Rstudio. Dengan menggunakan bantuan software Rstudio, hasil penerapan yang ditampilkan yaitu plot sebaran *cluster* dan jumlah anggota *cluster*. Hasil sebaran dari setiap *cluster* dari data pasien ditampilkan dalam bentuk plot yang diproses menggunakan Rstudio seperti gambar-gambar berikut ini :

```
K-means clustering with 2 clusters of sizes 32, 53
Cluster means:
  leko  trombo    hb    hct
1 6684.375 180437.50 14.43125 40.71875
2 5607.547  78396.23 14.22830 40.74906
```

Gambar 4.1 Hasil *K-Means 2 Cluster*

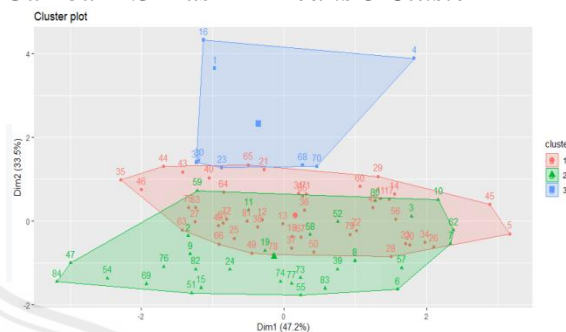
Gambar 4.2 Hasil Plot *K-Means 2 Cluster*

Pada hasil 2 *cluster* didapatkan anggota di masing-masing *cluster* yaitu 32 anggota



untuk *cluster 1* dan 53 anggota untuk *cluster 2*.

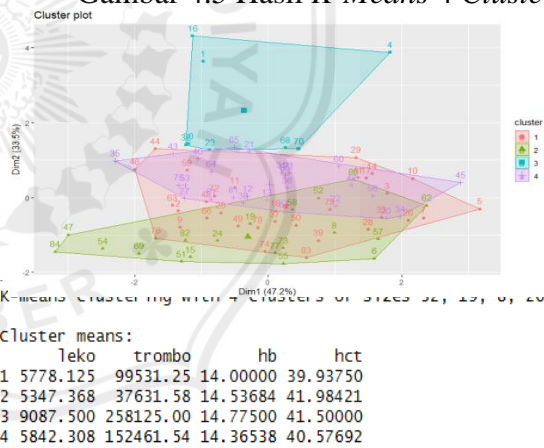
Gambar 4.3 Hasil *K-Means 3 Cluster*



Gambar 4.4 Hasil Plot *K-Means 3 Cluster*

Pada hasil 3 *cluster* didapatkan anggota di masing-masing *cluster* yaitu 47 anggota untuk *cluster 1*, 30 anggota untuk *cluster 2* dan 8 anggota untuk *cluster 3*.

Gambar 4.5 Hasil *K-Means 4 Cluster*



Gambar 4.6 Hasil Plot *K-Means 4 Cluster*

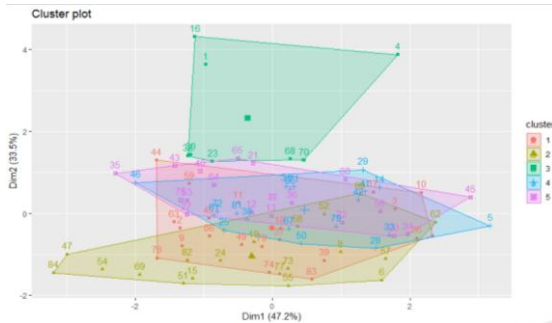
Pada hasil 4 *cluster* didapatkan anggota di masing-masing *cluster* yaitu 32 anggota untuk *cluster 1*, 19 anggota untuk *cluster 2*, 8 anggota untuk *cluster 3* dan 26 anggota untuk *cluster 4*.

K-means clustering with 5 clusters of sizes 21, 19, 8, 19, 18

```

Cluster means:
leuko trombo hb hct
1 5733.333 90523.81 14.30476 40.80952
2 5347.368 37631.58 14.53684 41.98421
3 9087.500 258125.00 14.77500 41.50000
4 5942.105 124052.63 13.84211 39.00000
5 5750.000 160611.11 14.33889 40.83333
    
```

Gambar 4.7 Hasil K-Means 5 Cluster



Gambar 4.8 Hasil Plot K-Means 5 Cluster

Pada hasil 5 cluster didapatkan anggota di masing-masing cluster yaitu 21 anggota untuk cluster 1, 19 anggota untuk cluster 2, 8 anggota untuk cluster 3, 19 anggota untuk cluster 4 dan 18 anggota untuk cluster 5

Dari hasil perhitungan elbow yang menentukan jumlah cluster terbaik ada di 3 cluster. Maka pada hasil berikut dari gambar 4.6 hasil cluster data pasien yang ada di 3 cluster, dimana terdapat hasil jumlah anggota di masing-masing cluster dan hasil plot. Pada titik berwarna merah merupakan anggota kelompok pertama, titik berwarna hijau merupakan anggota kelompok kedua dan titik berwarna biru merupakan anggota kelompok ketiga. Hasil yang diperoleh dari masing-masing anggota kelompok tersebut dapat dilihat di tabel 4.4, 4.5, dan 4.6.

Tabel 4.2 Anggota Cluster 1 Data Pasien DB

NO	Cluster 1				Leukosit
	Leukosit	Trombosit	Hemoglobin	Hematokrit	
5	4500	116000	10,2	32	Trombosit
...	...	...	...	...	
26	4300	96000	11,6	35	Hemoglobin
...	...	...	...	...	
28	2200	124000	12,8	36	Hematokrit
...	...	...	...	...	
35	7400	182000	17	48	...
...	...	...	...	...	...
44	11900	103000	15,9	46	...
...	...	...	...	...	...

Tabel 4.3 Anggota Cluster 2 Data Pasien DB

NO	Cluster 2			
	Leukosit	Trombosit	Hemoglobin	Hematokrit
2	6500	79000	15,5	46
...	...	...	...	...
10	8700	91000	11,8	33
...	...	...	...	...
47	6700	19000	17,8	50
51	3800	15000	16,2	44
...	...	...	...	...
62	7600	49000	11,7	32
...	...	...	...	...
80	10700	45000	12,6	36
...	...	...	...	...
83	1400	77000	13,9	39
84	5000	30000	16	57

Tabel 4.4 Anggota Cluster 3 Data Pasien DB

NO	Cluster 3			
	Leukosit	Trombosit	Hemoglobin	Hematokrit
1	10200	358000	15	45
4	12800	306000	11,8	35
16	18300	224000	15,3	43
23	7100	212000	15,2	44
30	49000	277000	16,5	43
...	...	...	...	...

Keterangan :

... : Data selanjutnya

Dari hasil tabel 4.2, 4.3, dan 4.4 untuk kolom yang berwarna jingga merupakan range atribut yang tertinggi dan untuk kolom berwarna kuning merupakan range atribut terendah. Berikut hasil range dari 3 cluster :

Tabel 4.5 Range Dari 3 Cluster

		C1	C2	C3
Leukosit	Terendah	2200	1400	4900
	Tertinggi	11900	10700	18300
Trombosit	Terendah	96000	15000	212000
	Tertinggi	182000	91000	358000
Hemoglobin	Terendah	10,2	11,7	11,8
	Tertinggi	17	17,8	16,5
Hematokrit	Terendah	32	32	35

	Tertinggi	48	57	45
--	-----------	----	----	----

4	37.440.462.614
5	30.220.312.131

Dari hasil tabel *range* diatas dapat disimpulkan dari setiap *cluster* memiliki nilai terendah dan tertinggi, maka untuk penjelasan setiap atribut dalam masing-masing *cluster* yaitu :

1. Untuk *cluster* 1 pada atribut leukosit, trombosit, hematokrit posisinya di antara *cluster* 2 dan 3. Selanjutnya atribut hemoglobin ada di posisi yang paling rendah dibandingkan *cluster* 2 dan 3.
2. Untuk *cluster* 2 pada atribut leukosit dan trombosit ada di posisi paling rendah dibandingkan *cluster* 1 dan 3. Selanjutnya atribut hemoglobin dan hematokrit ada di posisi paling tinggi dibandingkan *cluster* 1 dan 3.
3. Untuk *cluster* 3 pada atribut leukosit dan trombosit ada di posisi paling tinggi dibandingkan *cluster* 1 dan 2, pada atribut hemoglobin posisinya di antara *cluster* 1 dan 2 dan selanjutnya atribut hematokrit ada di posisi paling rendah dibandingkan *cluster* 1 dan 2.

### Implementasi Metode *Elbow*

Pada tahap pengujian ini mencari jumlah *cluster* terbaik dari data yang akan dikelompokkan, penentuan jumlah *cluster* dihitung dengan metode *Elbow*. Untuk masing-masing hasil *elbow* di setiap *cluster* dapat dilihat pada tabel 4.6 di bawah ini :

Tabel 4.6 Hasil SSE dari setiap *cluster*

Cluster	Hasil SSE
2	151.220.715.480
3	66.968.404.465

Setelah didapatkan hasil SSE dari setiap *cluster*, selanjutnya menghitung selisih SSE untuk mendapatkan nilai *elbow* sebagai penentuan berapa jumlah kelompok terbaik.

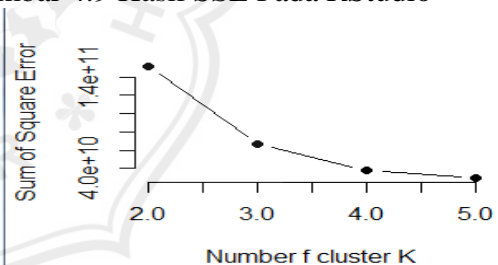
Tabel 4.7 Hasil Selisih SSE Dari Setiap *Cluster*

Cluster	SSE	Hasil Selisih
2 - 3	151.220.715.480 -	84.252.311.014
3 - 4	84.252.311.014 -	46.811.848.400
4 - 5	46.811.848.400 -	16.591.536.269

Dari tabel di atas dapat dilihat hasil selisih tertinggi ada di 3 *cluster*, maka dapat di tentukan *cluster* terbaik dalam penelitian ini berjumlah 3 *cluster*. Untuk implementasi metode *elbow* pada RStudio dapat dilihat pada gambar 4.9 dan gambar 4.10.

```
> sse
[1] 151220715480 66968404465 37440462614 30220312131
```

Gambar 4.9 Hasil SSE Pada RStudio



Gambar 4.10 Penerapan Metode *Elbow*

## VI. Kesimpulan dan Saran

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa :

1. Didapatkan hasil *cluster* terbaik ada pada 3 *cluster* yang dihitung menggunakan *elbow performance* yaitu 66.968.404.465.

2. Dari pengelompokkan metode *K-Means* dihasilkan 3 *cluster* yaitu :
  - a. *Cluster* 1 terdapat 47 data pasien
  - b. *Cluster* 2 terdapat 30 data pasien
  - c. *Cluster* 3 terdapat 8 data pasien

Beberapa saran yang dapat dijadikan dalam mengembangkan penelitian ini adalah:

1. Pengembangan selanjutnya disarankan dapat membangun sebuah sistem/aplikasi yang dapat lebih membantu untuk proses pendagnosisan pasien, karena pada penelitian ini hanya sebagai penerapan dari algoritma *K-Means*.
2. Untuk penentuan jumlah *cluster* terbaik bisa menggunakan metode yang lainnya tidak hanya menggunakan metode *elbow*.

### Daftar Pustaka

- WHO. 2011. *Haemoglobin Concentration for the Diagnosis of Anaemia and Assessment of Severity*. World Health Organization.
- WHO. 2009. *DENGUE :Guidelines for Diagnosis, Treatment, Prevention and Control*. World Health Organization.
- Ardilla, Y., Tjandrasa, H. Dan Arieshanti, I. 2014. "Deteksi Penyakit Epilepsi dengan Menggunakan Entropi Permutasi, K-Means Clustering dan Multilayer Perceptron," *Jurnal Teknik POMITS*, vol. 3, no. 1, pp. A70 – A74.
- Genis, G. 2008. Demam Berdarah. Yogyakarta : PT Bentang Pustaka.
- Kurane, I. 2007 Dengue Hemorrhagic Fever with Spesial Emphasis on Immunopathogenesis. Comparative Immunology, Microbiology & Infectious Disease. Vol 30:329-40.
- Chuansumrit A, Tangnaratchakit K. 2006. Pathophysiology and Management of Dengue Hemorrhagic Fever. Bangkok: Department of Pediatrics, Faculty of Medicine, Ramathibodi Hospital, Mahidol University.
- Hadinegoro, Rezeki S, Soegianto S, Soeroso T, Waryadi S. 2001. Tata Laksana Demam Berdarah Dengue di Indonesia. Jakarta: Ditjen PPM&PL Depkes&Kesos R.I.
- Ong J. O. 2013. "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University," *Jurnal Ilmiah Teknik Industri*, vol. 12, no. 1, pp. 10 - 20.
- Prilianti, K. R. dan Wijaya, H. 2014. "Aplikasi Text Mining Untuk Automasi Penentuan Tren Topik Skripsi Dengan Metode K-Means Clustering," *Jurnal Cybermatika*, vol. 2, no. 1, pp. 1-6.
- Windarto, A. P. 2017. "Penerapan Data Mining PadaEksporBuah-BuahanMenurut Negara TujuanMenggunakan K-Means Clustering," *Techno.COM*, vol. 16, no. 4, pp. 348–357.
- Agustina, S., Yhudo, D., Santoso, H., Marnasusanto, N., Tirtana, A. dan Khusnu, F. 2012. "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means," Universitas Brawijaya Malang, Malang.
- Siyamto, Y. 2017. "Pemanfaatan Data Mining DenganMetode Clustering UntukEvaluasiBiayaDokumenEkspor," vol. 1, no. 2, pp. 28–31.
- Madhulatha, T.S., 2012. *An Overview On Clustering Methods*. IOSR Journal ofEngineering, II(4), pp.719-725.
- Kodinariya, T. M. & Makwana, P. R., (2013). *Review on determining number of clusterin K-Means Clustering*. International Journal of Advance Research in Computer Science and Management Studies, I(6), pp. 90-95
- Bholowalia, P. & Kumar, A., 2014. *EBK-Means: A Clustering Techiniques based on Elbow Method and K-Means in WSN*.International Journal of Computer Application (0975-8887), IX(105), pp. 17-24.
- Irwanto, Purwananto, Y., dan Soelaiman, R., (2012). *Optimasi Kinerja Algoritma Klasterisasi K-Means untuk kuantisasi*

*Warna Citra.* Jurnal Teknik ITS, I(1),  
pp.197-202.

