

**CLUSTERING FAKTOR STRES PADA MAHASISWA AKTIF MENGGUNAKAN ALGORITMA K-MEANS DAN K-MODES**

**Jeffry Wijaya<sup>1</sup>, Trisha Magdalena Adelheid Januaviani<sup>2</sup>**

Universitas Matana, Banten, Indonesia

Email: jeffry.wijaya@student.matanauniversity.ac.id

---

**KATA KUNCI**

*Clustering, College Students, K-Means Algorithm, K-Modes Algorithm, Stress Factor*

**ABSTRACT**

*Stress is a part of human life that cannot be avoided. Stress is the biggest health problem experienced by all people, one of which is college students. Stress is influenced by external and internal factors. Some factors that cause stress to college students are environment, academics, financial, family, friendships, romantic relationships, health, self-confidence, career, hobbies/activities, and others. Using the K-Means Algorithm method, these factors can be divided into various clusters that have similar characteristics as triggers for stress to college students. Clustering is done using the R programming language, which shows the results of: There are 3 clusters using the K-Means algorithm, which: cluster 1 has 103 data (41.036%), cluster 2 has 72 data (28.685%), and cluster 3 has 76 data (30.279%). Cluster 1 has the most respondents compared to other clusters with the characteristics of respondents who experience stress due to academic, individual, family, and financial factors. This study found 4 clusters when using the K-Modes algorithm, which: cluster 1 has 92 data (36.65%), cluster 2 has 59 data (23.51%), and cluster 3 has 13 data (5.18%), and cluster 4 has 87 data (34.66%). Cluster 3 has the least number of respondents compared to other clusters with the characteristics of respondents who experience stress from individual and social factors.*

**ABSTRAK**

Stres merupakan bagian dari kehidupan manusia yang tidak bisa dihindari. Stres adalah permasalahan kesehatan terbesar yang dialami seluruh kalangan, salah satunya adalah mahasiswa. Stres dipengaruhi dari faktor eksternal maupun internal. Beberapa faktor yang menyebabkan terjadinya stres pada mahasiswa, diantaranya: lingkungan, akademis, finansial, keluarga, pertemanan, percintaan, kesehatan, kepercayaan diri, karir, kegiatan/hobi, dan lain-lain. Menggunakan metode Algoritma K-Means dan K-Modes, faktor-faktor tersebut dapat dibagi menjadi berbagai kluster yang memiliki karakteristik serupa sebagai pemicu terjadinya stres pada mahasiswa. *Clustering* dilakukan menggunakan bahasa pemrograman R yang menunjukkan hasil: Terdapat 3 *cluster* menggunakan algoritma K-Means, diantaranya: *cluster* 1 memiliki 103 data (41,036%), *cluster* 2 memiliki 72 data (28,685%),

dan *cluster* 3 memiliki 76 data (30,279%). *Cluster* 1 memiliki responden terbanyak dibandingkan dengan *cluster* lainnya dengan karakteristik responden yang mengalami stres akibat faktor perkuliahan, individual, keluarga, dan finansial. Penelitian ini mendapatkan 4 *cluster* ketika menggunakan algoritma K-Modes, diantaranya: *cluster* 1 memiliki 92 data (36,65%), *cluster* 2 memiliki 59 data (23,51%), dan *cluster* 3 memiliki 13 data (5,18%), dan *cluster* 4 memiliki 87 data (34,66%). *Cluster* 3 memiliki responden tersedikit dibandingkan dengan *cluster* lainnya dengan karakteristik responden yang mengalami stres dari faktor individual dan sosial.

---

## **PENDAHULUAN**

Mahasiswa merupakan seorang yang sedang menempuh pendidikan tinggi yang umumnya berkisar dari umur 18 sampai 25 tahun. Ketika masa perkuliahan, mahasiswa melakukan kegiatan akademik maupun non-akademik. Kegiatan akademik merupakan kegiatan yang berkaitan langsung dengan hal akademis seperti belajar, kerja kelompok, penelitian, tugas, dan lain-lain, sedangkan kegiatan non-akademik merupakan kegiatan-kegiatan di luar dari kewajiban akademis seperti organisasi, seni, olahraga, ekstrakurikuler, magang, dan sebagainya. Mahasiswa pada perguruan tinggi memiliki resiko yang tinggi terkena stres dari berbagai faktor atau stresor (Andiarna & Kusumawati, 2020).

Saat masa perkuliahan, mahasiswa dapat mengalami stres dari banyak hal seperti diri mahasiswa itu sendiri, lingkungan belajar, keluarga, dan lain-lain (Lestari, 2021). Hal tersebut memicu terbentuknya tekanan-tekanan yang dihadapi oleh mahasiswa sehingga menjadi faktor terjadinya stres. Tekanan yang diberikan bisa merupakan harus mampu belajar mandiri, tugas menumpuk, omongan orang tua, membengkaknya biaya kuota internet untuk belajar daring, dan sebagainya (Andiarna & Kusumawati, 2020). Tanggungjawab dan tuntutan secara akademik dan banyaknya masalah yang dihadapi mahasiswa dapat menjadi bagian stres pada mahasiswa itu sendiri (Hamzah & Hamzah, 2020). Stres merupakan salah satu gangguan kesehatan mental yang tidak dapat dihindari dan dapat dialami oleh siapapun. Adapun sumber-sumber penyebab stres, diantaranya: internal (pola pikir, kepribadian, dan keyakinan) dan eksternal (beban pelajaran, tekanan untuk mendapatkan nilai bagus, dorongan meningkatkan status sosial, dan tekanan dari orang tua) (Siregar & Putri, 2020). Faktor eksternal penyebab stres diantaranya: lingkungan, akademis, finansial, keluarga, pertemanan, percintaan, karir/pekerjaan, kegiatan/hobi, dan lain-lain, sedangkan faktor internal yang menyebabkan stres adalah perilaku, kesehatan, kepercayaan diri, dan lain-lain (Putu, Novita, Bagus, & Dwidsamara, 2021). Stres sendiri merupakan hal yang tidak bisa dianggap remeh karena dapat memicu korbannya untuk melakukan hal yang berdampak buruk bagi masyarakat maupun dirinya sendiri seperti mudah sakit, tidak bersemangat atau mudah menyerah bahkan bunuh diri (Saefudin, 2019).

Beberapa faktor terjadinya stres dapat di-*cluster* menjadi sebuah kelompok yang memiliki karakteristik serupa. *Clustering* merupakan teknik optimasi tujuan sebuah objek secara global sehingga didapatkan suatu kriteria yang memiliki kemiripan dalam sebuah pola (Golalipour, Akbari, Hamidi, Lee, & Enayatifar, 2021). *Clustering* dapat dilakukan dengan berbagai algoritma, diantaranya: *K-Means* dan *K-Modes*. Algoritma *K-Means* merupakan algoritma yang digunakan untuk melakukan *clustering* dengan menggunakan jarak antar data dengan centroid (El-Mandouh, Abd-Elmegid, Mahmoud, & Haggag, 2019). *K-Means* juga

merupakan metode yang cukup mudah namun cocok untuk melakukan *clustering* dengan tepat (El-Mandouh et al., 2019). Algoritma *K-Modes* merupakan algoritma *K-Means* yang dikembangkan oleh Huang pada tahun 1998 untuk melakukan *clustering* data kategorikal. Menggunakan skala 0 dan 1, metode ini merupakan metode yang lebih cepat dibandingkan *K-Means* karena melakukan iterasi lebih sedikit untuk mendapatkan hasil optimal (Sajidha, Chodnekar, & Desikan, 2021).

## **METODE PENELITIAN**

Penelitian yang digunakan untuk penelitian ini dikumpulkan menggunakan Google Forms berbentuk kuesioner berisi 3 variabel identitas dan 77 variabel alat ukur dengan 251 baris atau responden (Jogiyanto Hartono, 2019). Teknik pengumpulan data pada penelitian ini menggunakan alat ukur bersifat kuantitatif yang menggunakan skala 5-point Likert-type scale mulai dari 1 sampai dengan 5. 5-point Likert-type scale adalah sebuah skala alat ukur yang bertujuan untuk mengkonversi suatu angka menjadi skala pernyataan yang sangat tidak menyenangkan, sangat tidak memuaskan, dan sangat tidak setuju sampai sangat menyenangkan, sangat memuaskan, dan sangat setuju terhadap suatu subjek atau butir (Tanujaya, Prahmana, & Mumu, 2022). Penilaian angka 1 berarti Sangat Tidak Setuju (STS) dan 5 berarti Sangat Setuju (SS). Data diperoleh menggunakan gabungan dari 5 alat ukur untuk masing-masing faktor stres, yaitu akademik, individu, sosial, keluarga, dan keuangan yang telah divalidasi.

Subjek penelitian ini adalah mahasiswa aktif di beberapa wilayah Indonesia khususnya Jabodetabek dan Surabaya pada universitas swasta atau negeri berumur 18 sampai 23 tahun (Thama, 2021). Perangkat atau aplikasi yang digunakan untuk mendukung penelitian ini adalah Microsoft Excel dan RStudio. Aplikasi RStudio merupakan software pengolahan data bersifat open source yang terintegrasi dengan bahasa pemrograman R dan Python. RStudio didukung langsung menggunakan package-package yang ada di dalam aplikasi agar dapat melakukan analisis data mulai dari bentuk sederhana hingga kompleks. Metode *clustering* yang digunakan pada penelitian ini ada 2, diantaranya algoritma *K-Means* untuk data numerik dan algoritma *K-Modes* untuk data kategorik (Wardhani & Khrisna, 2016).

## **HASIL DAN PEMBAHASAN**

Pembersihan data adalah proses pengolahan data yang dilakukan untuk memastikan apakah data yang digunakan sudah tidak data kosong, duplikat, dan outliers.

### **Stuktur Data**

Sebelum melakukan pembersihan data, tipe data tiap variabel perlu diketahui. Variabel  $j_i$  merupakan variabel pertanyaan dimana,  $j$  = faktor stres kuesioner dan  $i$  = pertanyaan ke- $i$ . Variabel  $j$  dibagi menjadi 5, yaitu: A = Academic, I = Individual, S = Social, Fa = Family, Fi = Finance. Struktur data pada data dapat dilihat pada Tabel 1.

**Tabel 1**  
**Struktur Data Survei**

<i>Variabel</i>	<i>Tipe Data</i>
Jenis Kelamin	chr
Umur	num
Universitas	chr
A01	num
A02	num
...	num
Fi13	num

Berdasarkan Tabel 1, dapat diketahui bahwa tipe data untuk setiap beberapa variabel merupakan num (data numerikal atau angka) atau chr (data karakter atau teks). Karena variabel umur pada data penelitian ini bukan merupakan variabel yang perlu dihitung maka tipe data variabel umur peneliti ubah agar menjadi chr atau karakter.

**Pengecekan data kosong**

Pengecekan data kosong atau data NA dilakukan untuk mencari data yang tidak memiliki sebuah nilai atau memiliki nilai tidak valid (Triscowati & Jayanti, 2021). Sebelum melakukan visualisasi dan pengolahan, data yang invalid seperti NA harus dihapus terlebih dahulu. Hasil data kosong dapat dilihat pada Gambar 1.



**Gambar 1**  
**Hasil Pengecekan Data Kosong**

Berdasarkan Gambar 1, diketahui data penelitian ini tidak memiliki data kosong atau NA.

**Pengecekan data duplikat**

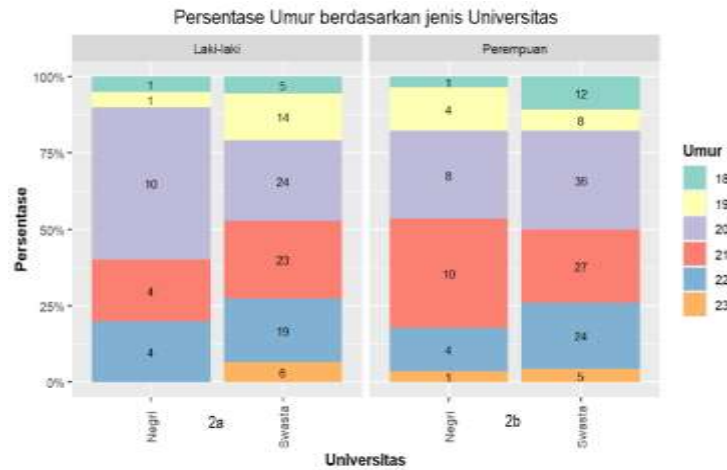
Pengecekan data duplikat dilakukan untuk mencari apakah ada 2 baris atau lebih data yang memiliki nilai sama persis. Proses ini dilakukan dengan cara membandingkan jumlah baris data dengan jumlah baris data yang unik menggunakan code pada RStudio sehingga didapatkan hasil bahwa jumlah data unik dengan jumlah data adalah sama, sehingga dapat disimpulkan tidak ada baris yang duplikat pada data penelitian ini.

**Pembersihan outliers**

Dikarenakan topik utama dalam penelitian ini merupakan stres, maka peneliti tidak melakukan pembersihan outliers dikarenakan adanya kemungkinan seorang mahasiswa memang mengalami stres pada banyak bidang maupun tidak merasa terbebani sama sekali.

**Visualisasi Data Penelitian**

Visualisasi data penelitian dilakukan untuk mempermudah pembaca memahami bentuk dan sebaran setiap variabel maupun pertanyaan dalam penelitian ini (Resnatika, Sukaesih, & Kurniasih, 2018). Tabel 3 merupakan sebaran data berdasarkan identitas sedangkan Gambar 2 merupakan beberapa hasil visualisasi data penelitian ini.



**Gambar 2**  
**Diagram Percent Stacked Barchart Data Keseluruhan**

Berdasarkan Gambar 2 dapat dilihat seluruh distribusi data untuk ketiga variabel identitas yang dimiliki. Grafik 2a mewakili jenis kelamin laki-laki sedangkan Grafik 2b mewakili jenis kelamin perempuan.

**Tabel 3**  
**Sebaran Data Berdasarkan Variabel Identitas**

Variable Identitas		n	Proporsi (%)
Jenis Kelamin	<b>Laki-laki</b>	111	44,22%
	<b>Perempuan</b>	140	55,78%
Umur	<b>18</b>	19	7,57%
	<b>19</b>	27	10,76%
	<b>20</b>	78	31,08%
	<b>21</b>	64	25,50%
	<b>22</b>	51	20,32%
	<b>23</b>	12	4,78%
Universitas	<b>Nagri</b>	48	19,12%
	<b>Swasta</b>	203	80,88%
Total		251	100,00%

### Mengubah Skor Data

Sebelum melakukan *clustering* menggunakan total skor dari masing-masing kuesioner, terdapat pertanyaan-pertanyaan yang bersifat positif sehingga skor pada pertanyaan-pertanyaan tersebut harus dibalik terlebih dahulu. Hal ini dikarenakan jawaban skor 1 (Sangat Tidak Setuju) menjadi skor 5. Pertanyaan-pertanyaan yang bersifat positif dapat dilihat pada Tabel 4.

**Tabel 4**  
**Daftar Pertanyaan Bersifat Positif**

101	Saya merasa beristirahat dengan cukup
107	Saya seringkali merasa melakukan hal yang saya sukai
110	Saya merasa tenang
113	Saya merasa penuh energi
117	Saya merasa aman dan terlindungi
121	Saya merasa tenang dengan diri saya sendiri
125	Saya adalah pribadi yang murah hati
129	Saya memiliki waktu yang cukup untuk diri sendiri
S02	Orang-orang di sekitar saya memberikan lebih banyak kebahagiaan dibandingkan masalah
S03	Saya memiliki teman-teman yang baik
S08	Kebanyakan orang mengagumi cara saya hidup
S09	Prestasi yang saya capai diakui dengan layak
S10	Saya tidak pernah dibiarkan sendirian saat menyelesaikan masalah
S11	Saya memiliki orang yang dapat diandalkan

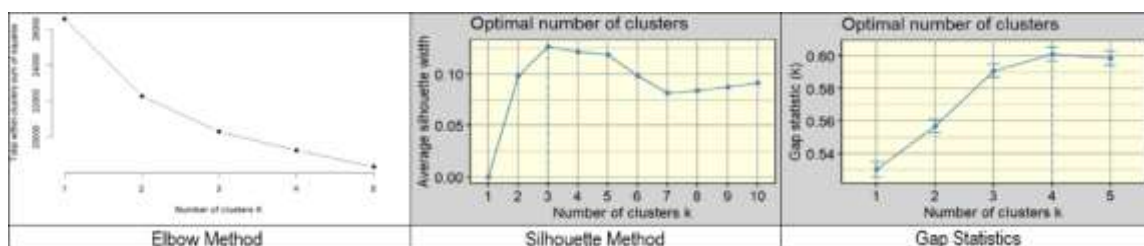
Setelah skor empat belas pertanyaan tersebut dibalik, proses *clustering* dapat dilakukan.

### K-Means Clustering

K-Means *clustering* dilakukan pada data numerik dengan cara mencari nilai  $k$  optimal (Refialy, Maitimu, & Pesulima, 2021). Setelah didapatkan  $k$  optimal untuk data penelitian ini, maka dilakukan *clustering* dengan aplikasi RStudio.

#### a. Mencari $k$ optimal

$k$  optimal dicari untuk menentukan jumlah *cluster* yang dibentuk oleh data penelitian (Nainggolan & Lumbantoruan, 2018). Ada berbagai cara untuk mencari  $k$  optimal untuk algoritma K-Means, diantaranya: *Elbow Method*, *Silhouette Method*, dan *Gap Statistic*.



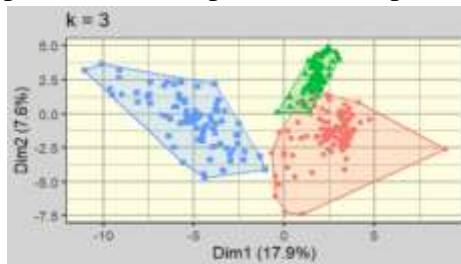
**Gambar 3**  
**Hasil k optimal Menggunakan 3 Metode di RStudio**

Berdasarkan Gambar 3 dapat diperkirakan nilai  $k$  optimal untuk *Elbow Method* adalah 2 atau 3 yang dinyatakan karena bentuk sikut mulai lurus pada  $k=2$  atau  $k=3$ , untuk *silhouette method* dapat diketahui bahwa nilai  $k$  optimalnya adalah 3 karena titik tertinggi (garis putus-putus) berada pada  $k=3$ , sedangkan menggunakan metode *Gap Statistic* dapat diketahui nilai  $k$  optimal adalah 4. Karena menggunakan code pada R terlihat garis putus-putus pada  $k=4$ . Dari

ketiga metode diatas, maka didapatkan bahwa k optimal untuk data penelitian ini adalah k=3 dari nilai bilangan bulat rata-rata ketiga metode tersebut.

b. Melakukan *clustering*

Setelah mendapatkan k optimal, maka dapat dilakukan proses *clustering*.



**Gambar 4**

**Plot Hasil K-Means Clustering dengan k=3**

Berdasarkan Gambar 4 dapat diketahui setiap titik bentuk mewakili sebuah data. *Cluster* 1 memiliki wilayah berwarna merah, *cluster* 2 memiliki wilayah berwarna hijau, dan *cluster* 3 memiliki wilayah berwarna biru.

**Tabel 5**  
**Tabel Distribusi Data K-Means Clustering per Cluster**

Variable Identitas		Cluster 1	Cluster 2	Cluster 3	Total	Proporsi (%)
Jenis Keliamin	Laki-laki	46	32	32	111	44,223%
	Perempuan	57	39	44	140	55,777%
Umur	18	10	5	4	19	7,570%
	19	13	5	9	27	10,757%
	20	27	23	28	78	31,076%
	21	29	17	18	64	25,498%
	22	17	20	14	51	20,319%
	23	7	2	3	12	4,781%
Universitas	Negri	37	0	11	48	19,124%
	Swasta	66	72	65	203	80,876%
Total		103	72	76	215	100,00%
Proporsi		41,036%	28,685%	30,279%	100,00%	

Berdasarkan Tabel 5 dapat kita ketahui bahwa *cluster* 1 memiliki 103 data (41,036%), *cluster* 2 memiliki 72 data (28,685%), dan *cluster* 3 memiliki 76 data (30,279%). *Cluster* 1 memiliki responden terbanyak dibandingkan dengan *cluster* lainnya.

**K-Modes Clustering**

K-Modes *clustering* merupakan metode *clustering* yang digunakan pada data kategorik, oleh karena itu sebelum mencari nilai k optimal, perlu dilakukan transformasi data dari numerik menjadi kategorik.

1. Mengubah bentuk data

Bentuk awal data penelitian ini merupakan data numerik berupa skor jawaban untuk setiap pertanyaan. Oleh karena itu, untuk mengubah jawaban tersebut menjadi kategorik, dilakukan tahapan sebagai berikut:

## Clustering Faktor Stres Pada Mahasiswa Aktif Menggunakan Algoritma K-Means Dan K-Modes

### a. Menormalisasi total skor

Normalisasi jawaban kuesioner dilakukan untuk mengubah total data menjadi skala 0-1 menggunakan Persamaan (14).

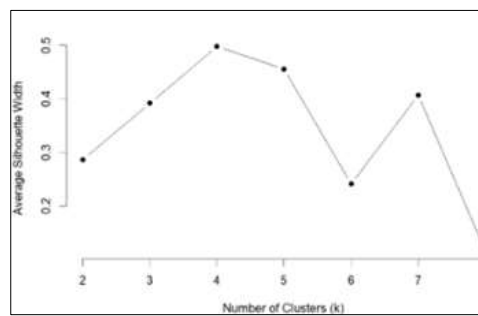
$$Skor_{Normalized} = \frac{Total_{Skor} - Total_{SkorMinimal}}{Total_{SkorMaximal}}$$

### b. Mengubah menjadi kategorik

Dikarenakan setiap kuesioner tidak memiliki tolak ukur pasti untuk membatasi tingkat stres berdasarkan skor, penelitian ini menggunakan nilai  $Skor_{Normalized} > 0,6$  dikategorikan sebagai stres karena semakin tinggi nilai skornya maka individu tersebut dinyatakan semakin stres (Levenstein et al., 1993).

## 2. Mencari k optimal

k optimal dicari untuk menentukan jumlah *cluster* yang dibentuk oleh data penelitian. Karena data merupakan data kategorik, maka metode untuk mencari k optimal sangat terbatas. Oleh karena itu digunakan Silhouette Method untuk mencari k optimal.



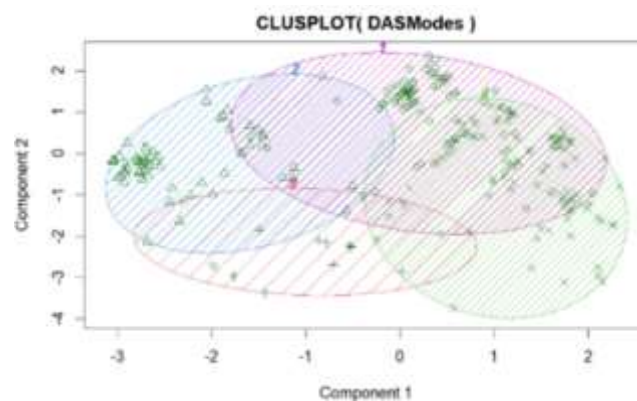
**Gambar 5**

**Hasil k optimal Silhouette Method di Rstudio**

Berdasarkan Gambar 5 dan cara menentukan Silhouette Method (melihat titik maksimum pada grafik) maka dapat didapatkan nilai k optimal adalah k=4.

## 3. Melakukan *clustering*

Setelah mendapatkan k optimal, maka dapat dilakukan proses *clustering*.



**Gambar 6**

**Plot Hasil K-Modes Clustering dengan k=4**



Berdasarkan Gambar 6 dapat diketahui setiap titik bentuk mewakili sebuah data. Cluster 1 memiliki wilayah berwarna magenta, cluster 2 memiliki wilayah berwarna biru, cluster 3 memiliki wilayah berwarna merah, dan cluster 4 memiliki wilayah berwarna hijau.

**Tabel 6**  
**Tabel Distribusi Data K-Means Clustering per Cluster**

Variable Identitas		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Total	Proporsi(%)
Jenis Kelamin	Laki-laki	41	27	3	40	111	44,223%
	Perempuan	51	32	10	47	140	55,777%
Umur	18	9	3	0	7	19	7,570%
	19	6	6	3	12	27	10,757%
	20	23	21	5	29	78	31,076%
	21	26	15	1	22	64	25,498%
	22	24	11	3	13	51	20,319%
	23	4	3	1	4	12	4,781%
Universitas	Negri	8	9	3	28	48	19,124%
	Swasta	84	50	10	59	203	80,876%
Total		92	59	13	87	251	100,00%
Proporsi (%)		36,653%	23,506%	5,179%	34,661%	100,00%	

Berdasarkan Tabel 6 dapat diketahui bahwa *cluster* 1 memiliki 92 data (36,65%), *cluster* 2 memiliki 59 data (23,51%), dan *cluster* 3 memiliki 13 data (5,18%), dan *cluster* 4 memiliki 87 data (34,66%). Menggunakan code di RStudio dengan data hasil *clustering* K-Modes dapat diperoleh hasil total “Yes” untuk setiap *cluster* seperti Tabel 7.

**Tabel 7**  
**Total “Yes” Untuk Masing-Masing Kuesioner Berdasarkan Cluster**

Cluster	Academic	Individual	Social	Family	Financial
1	91	16	5	15	89
2	7	4	1	2	4
3	4	13	11	1	3
4	82	70	19	73	84

Berdasarkan Tabel 7, dapat disimpulkan bahwa *cluster* 1 menunjukkan adanya kumpulan responden yang mengalami stres karena faktor akademik, dan keuangan, *cluster* 2 menunjukkan kumpulan responden yang hampir tidak mengalami stres dari faktor manapun, *cluster* 3 menunjukkan kumpulan responden yang mengalami stres karena faktor individual (internal) dan sosial, sedangkan *cluster* 4 menunjukkan kumpulan responden yang mengalami stres dari semua faktor selain faktor sosial.

## KESIMPULAN

*Cluster* 1 merupakan *cluster* yang memiliki responden paling banyak. *Cluster* ini terdiri atas 92 responden. Mayoritas responden pada *cluster* ini berumur 20 sampai 22 tahun. Responden pada *cluster* ini memiliki persentase mahasiswa di universitas negeri paling

sedikit dibandingkan dengan *cluster* lainnya dengan nilai 8,696%. Penyebab stres utama pada *cluster* ini adalah faktor akademik dan finansial. *Cluster* ini memiliki faktor stres yang sama dengan *cluster* 2 pada algoritma K-Means. *Cluster* 2 merupakan *cluster* yang memiliki jumlah responden menengah. *Cluster* ini terdiri atas 59 responden. Mayoritas responden pada *cluster* ini berumur 20 tahun dengan jumlah 21 responden. Diketahui tidak ada penyebab stres utama pada *cluster* ini untuk faktor-faktor yang dipertanyakan. *Cluster* ini memiliki karakteristik yang sama dengan *cluster* 3 pada algoritma K-Means.

*Cluster* 3 merupakan *cluster* yang memiliki responden paling sedikit. *Cluster* ini terdiri atas 13 responden. Mayoritas responden pada *cluster* ini berumur 20 tahun dengan jumlah 5 responden. Pada *cluster* ini tidak ada mahasiswa berumur 18 tahun. Penyebab stres pada *cluster* ini adalah faktor individual dan sosial. *Cluster* ini merupakan *cluster* baru yang tidak didapatkan ketika menggunakan  $k=3$  pada algoritma K-Means. Seratus persen *cluster* ini mengalami stres secara individual. *Cluster* 4 merupakan *cluster* yang memiliki jumlah responden cukup banyak. *Cluster* ini terdiri atas 87 responden. Mayoritas responden pada *cluster* ini berumur 20 dan 21 tahun. Penyebab stres utama pada *cluster* ini adalah faktor akademik, individual, keluarga, dan finansial. *Cluster* ini memiliki karakteristik yang sama dengan *cluster* 1 pada algoritma K-Means.

## DAFTAR PUSTAKA

- Andiarna, Funsu, & Kusumawati, Estri. (2020). Pengaruh pembelajaran daring terhadap stres akademik mahasiswa selama pandemi Covid-19. *Jurnal Psikologi*, 16(2), 139–149. [Google Scholar](#)
- El-Mandouh, Amira M., Abd-Elmegid, Laila A., Mahmoud, Hamdi A., & Haggag, Mohamed H. (2019). Optimized K-means clustering model based on gap statistic. *International Journal of Advanced Computer Science and Applications*, 10(1). [Google Scholar](#)
- Golalipour, Keyvan, Akbari, Ebrahim, Hamidi, Seyed Saeed, Lee, Malrey, & Enayatifar, Rasul. (2021). From clustering to clustering ensemble selection: A review. *Engineering Applications of Artificial Intelligence*, 104, 104388. [Google Scholar](#)
- Hamzah, B., & Hamzah, Rahmawati. (2020). Faktor-faktor yang berhubungan dengan tingkat stres akademik pada mahasiswa stikes graha medika. *Indonesian Journal for Health Sciences*, 4(2), 59–67. [Google Scholar](#)
- Jogiyanto Hartono, M. (2019). *Strategi Penelitian Bisnis*. Penerbit Andi. [Google Scholar](#)
- Lestari, Ani. (2021). Strategi Mahasiswa dalam Meningkatkan Minat Belajar, Self-Efficacy, Self Regulated pada Pembelajaran Daring di Masa Pandemi Covid-19. *Didaktis: Jurnal Pendidikan Dan Ilmu Pengetahuan*, 21(3). [Google Scholar](#)
- Nainggolan, Rena, & Lumbantoruan, Gortap. (2018). Optimasi performa cluster K-Means menggunakan Sum of Squared Error (SSE). *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, 2(2), 103–108. [Google Scholar](#)
- Putu, N., Novita, M., Bagus, I., & Dwidsamara, G. (2021). Implementation of K-Modes Algorithm for Clustering of Stress Causes in University Students. *J. Elektron. Ilmu Komput. Udayana*, 9(3), 419–428. [Google Scholar](#)
- Refialy, Leonardo Petra, Maitimu, Hervin, & Pesulima, Melgusen Soyano. (2021). Perbaikan kinerja clustering k-means pada data ekonomi nelayan dengan perhitungan sum of square error (sse) dan optimasi nilai k cluster. *Techno. Com*, 20(2), 321–329. [Google Scholar](#)

*Clustering Faktor Stres Pada Mahasiswa Aktif Menggunakan Algoritma K-Means Dan K-Modes*

- Resnatika, Ajeng, Sukaesih, Sukaesih, & Kurniasih, Nuning. (2018). Peran infografis sebagai media promosi dalam pemanfaatan perpustakaan. *Jurnal Kajian Informasi & Perpustakaan*, 6(2), 183–196. [Google Scholar](#)
- Saefudin, Wahyu. (2019). *Mengembalikan Fungsi Keluarga*. Ide Publishing. [Google Scholar](#)
- Sajidha, S. A., Chodnekar, Siddha Prabhu, & Desikan, Kalyani. (2021). Initial seed selection for K-modes clustering—a distance and density based approach. *Journal of King Saud University-Computer and Information Sciences*, 33(6), 693–701. [Google Scholar](#)
- Siregar, Ilham Khairi, & Putri, Sefni Rama. (2020). Hubungan self-efficacy dan stres akademik mahasiswa. *Consilium: Berkala Kajian Konseling Dan Ilmu Keagamaan*, 6(2), 91–95. [Google Scholar](#)
- Tanujaya, Benidiktus, Prahmana, Rully Charitas Indra, & Mumu, Jeinne. (2022). *Likert scale in social sciences research: Problems and difficulties*. [Google Scholar](#)
- Thama, Anggatra Dewi. (2021). *Analisis Mobilitas Mahasiswa Asal Kota–Kota Besar Di Pulau Jawa Menuju Kota Surakarta*. Universitas Muhammadiyah Surakarta. [Google Scholar](#)
- Triscowati, Dwi Wahyu, & Jayanti, Lia Dwi. (2021). Penilaian Kredit Pada Data Tak Seimbang Menggunakan Random Forest: Credit Scoring In Imbalance Data Using Random Forest. *Jurnal Ilmiah Komputasi Dan Statistika*, 1(1), 25–31. [Google Scholar](#)
- Wardhani, Anindya Khrisna, & Khrisna, Anindya. (2016). Implementasi Algoritma K-Means untuk pengelompokkan penyakit pasien pada puskesmas kajeun pekalongan. *J. Transform*, 14(1), 30–37. [Google Scholar](#)