

BAB 4

HASIL DAN PEMBAHASAN

4.1 Pengumpulan Data

Data yang digunakan mahasiswa FST UINSU Medan melalui pengisian kuesioner. Data mahasiswa yang digunakan usia awal merokok (x_1), jumlah rokok (bungkus) (x_2), prodi (x_3), jenis rokok (x_4) dan faktor merokok (x_5). Pelabelan untuk prodi yaitu (1) Ilmu Komputer, (2) Sistem Informasi, (3) Matematika, (4) Biologi dan (5) Fisika. Pelabelan untuk jenis rokok yaitu (1) Marlboro, (2) Sampoerna A Mild dan (3) Gudang Garam. Pelabelan untuk faktor merokok yaitu (1) Keluarga, (2) Teman dan (3) Media Massa (Iklan Rokok).

Tabel 4.1 : Data Awal

No	Nama Mahasiswa	Usia Awal Merokok	Jumlah Rokok (Bungkus)	Prodi	Jenis Rokok	Faktor Merokok
1	Fauzan	15	2	2	3	2
2	Abdul	17	1	2	3	1
3	Rizki	13	2	1	1	1
4	Rafly	11	2	3	3	2
5	Iwal	14	1	2	1	3
6	Fauzan	13	2	2	3	2
7	Jabal	16	3	1	2	1
8	Pradana	14	2	3	2	1
9	Shafwan	17	1	3	3	2
10	Amri	14	2	1	3	3

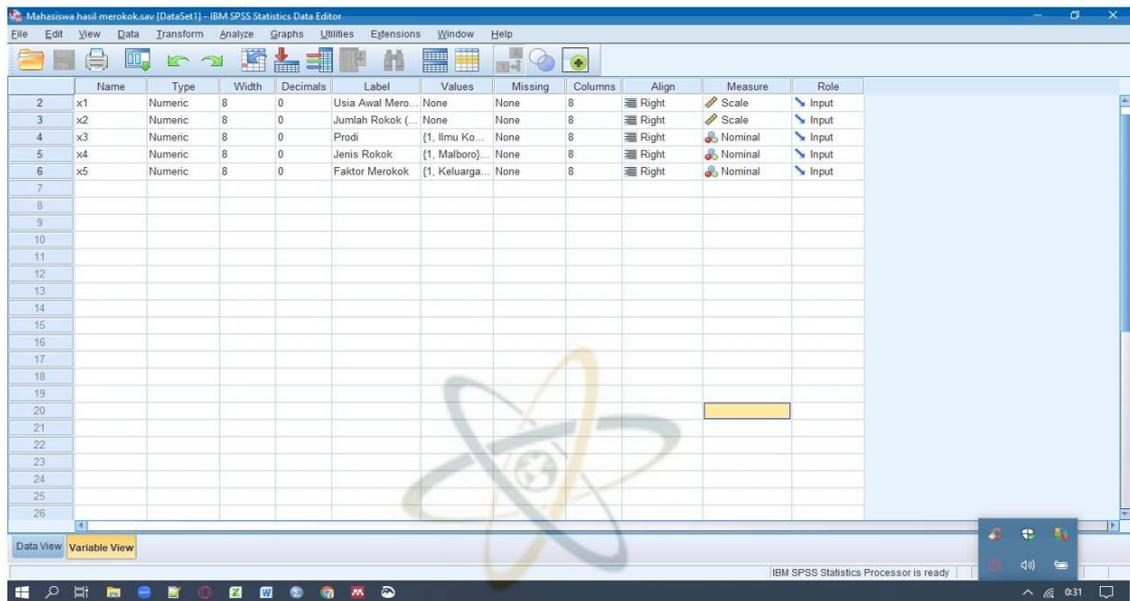
dan seterusnya sampai dengan 65 responden.

4.2 Asumsi Non Multikolinearitas

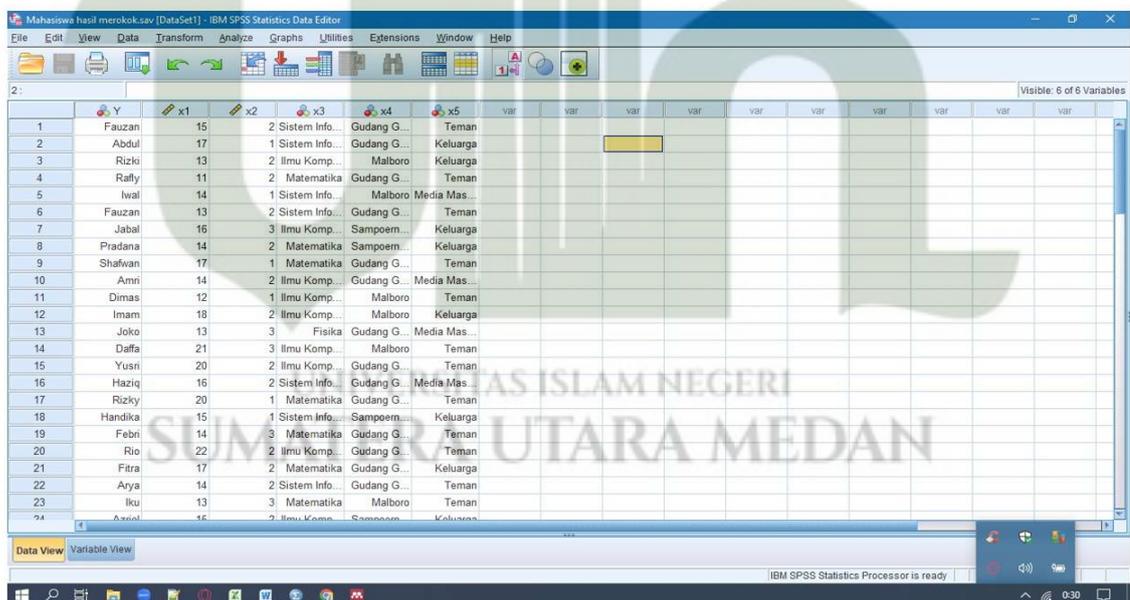
Pada tabel diatas dapat diketahui kuesioner mahasiswa merokok di FST UINSU Medan. Bila tidak ada korelasi signifikan antara variabel penjelas (independen), hal ini dikenal sebagai asumsi multikolinearitas. Program SPSS dapat digunakan untuk menentukan nilai VIF.

Cara menghitung nilai VIF:

1. Masuk ke SPSS isi *variable view* kemudian *Data View*



Gambar 4.1 *Variable View*



Gambar 4.2 *Data View*

2. Klik *Analyze - Regression - Linear* - Nama Mahasiswa (*dependent*) - x_1, x_2, x_3, x_4, x_5 (*Independent*) - *Statistics - collinearity diagnostics* setelah itu klik *continue* dan ok.

Tabel 4.2 menampilkan temuan uji multikolinearitas yang dilakukan dengan SPSS.

Tabel 4.2 : Hasil Uji Asumsi Non-Multikolinearitas

Variabel	VIF
x_1	1,041
x_2	1,070
x_3	1,089
x_4	1,033
x_5	1,026

Melihat nilai VIF lebih kecil dari 10,00 maka tidak terjadi multikolinearitas, sehingga asumsi non multikolinearitas terpenuhi.

4.3 Proses *Clustering*

Proses *clustering* menggunakan algoritma *K-Prototypes* untuk mengelompokkan data mahasiswa kecanduan merokok di FST UINSU Medan dengan $k = 3$.

1. Temukan jumlah klaster (k) yang perlu dikembangkan; misalnya, $k = 3$.
2. Sebagai *centroid* awal, pilih tiga item secara acak: objek ke-46, ke-51, dan ke-59.

Tabel 4.3 : *Centroid* Pertama Pengelompokan dengan $k=3$

Objek ke-i	<i>Centroid</i>	Indeks Variabel ke- (p)				
		1	2	3	4	5
46	c_1	13	2	1	3	3
51	c_2	12	1	2	2	1
59	c_3	15	3	3	1	2

Menentukan standar deviasi menggunakan persamaan 2.9. Sebelum mencari hasil standar deviasi mencari nilai mean, median, minimum dan maximum.

1. Mean

$$\begin{aligned}
 \bar{x} &= \frac{x_1+x_2+\dots+x_{65}}{n} \\
 &= \frac{15+17+13+11+14+13+16+14+17+14+12+18+13+21}{65} \\
 &\quad + \frac{20+16+20+15+14+22+17+14+13+15+18+15+20+}{65} \\
 &\quad + \frac{18+17+17+15+12+16+14+13+15+19+22+16+11+13}{65} \\
 &\quad + \frac{21+21+19+11+13+20+18+21+23+12+22+14+17+}{65} \\
 &\quad + \frac{21+22+20+20+15+16+22+18+22+12+21}{65} \\
 \frac{1086}{65} &= 16,707 = 16,71
 \end{aligned}$$

$$(x_i - \bar{x})^2 = (11 - 16,71)^2 = 32,604$$

$$(x_i - \bar{x})^2 = (14 - 16,71)^2 = 7,344$$

Dan seterusnya sampai ke 65 responden

Penjumlahan data ke 1 sampai ke 65 untuk $(x_i - \bar{x})^2$ adalah 759,38.

Standar deviasi menggunakan persamaan 2.9:

$$s^2 = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{759,38}{65-1}} = \sqrt{\frac{759,38}{64}} = 11,865$$

$$s = \sqrt{s^2} = \sqrt{11,865} = 3,445$$

Untuk x_2 :

$$(x_i - \bar{x})^2 = (2 - 1,75)^2 = 0,062$$

$$(x_i - \bar{x})^2 = (1 - 1,75)^2 = 0,562$$

$$(x_i - \bar{x})^2 = (2 - 1,75)^2 = 0,062$$

$$(x_i - \bar{x})^2 = (2 - 1,75)^2 = 0,062$$

$$(x_i - \bar{x})^2 = (1 - 1,75)^2 = 0,562$$

Dan seterusnya sampai ke 65 responden

Penjumlahan data ke 1 sampai ke 65 untuk $(x_i - \bar{x})^2$ adalah 36,03.

Standar deviasi menggunakan persamaan 2.9:

$$s^2 = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{36,03}{65-1}} = \sqrt{\frac{36,03}{64}} = 0,563$$

$$s = \sqrt{s^2} = \sqrt{0,563} = 0,751$$

Hasil standar deviasi menggunakan SPSS adalah:

1. Masuk ke SPSS isi *variable view* kemudian data view
2. Klik *Analyze - Descriptive Statistics - Frequencies* - Masukkan variabel numerik x_1 dan x_2 ke dalam *variable(s) statistics - mean - median - minimum - maximum - Std. Deviation* dan klik *continue*.

Tabel 4.4 : Standar Deviasi

Statistics			
		X_1	X_2
N	Valid	65	65
	Missing	0	0
Mean		16,71	1,75
Median		16,00	2,00
Std. Deviation		3,445	0,751
Minimum		11	1
Maximum		23	3

Diketahui bahwa $\sigma_1 = 3,445$ dan $\sigma_2 = 0,751$ sehingga nilai koefisien gamma (γ) adalah sebagai berikut:

Menentukan koefisien gamma (γ) menggunakan persamaan 2.8.

$$\gamma = \frac{1}{2} (3,445 + 0,751) = \frac{1}{2} (4,196) = 2,098$$

Selanjutnya, gunakan ukuran jarak campuran dalam persamaan 2.7 untuk mendapatkan jarak antara semua objek data dan *centroid* awal.

Misalnya, rumus berikut dapat digunakan untuk menentukan jarak antara objek pertama dan masing-masing *centroid* kluster pertama (c_1), *centroid* kluster kedua (c_2), dan *centroid* kluster ketiga (c_3).

$$\begin{aligned} d(x_1, c_1) &= ((x_{11} - c_{11})^2 + (x_{12} - c_{12})^2) + \gamma (\delta(x_{13}, c_{13}) + \delta(x_{14}, c_{14}) + \delta(x_{15}, c_{15})) \\ &= ((15 - 13)^2 + (2 - 2)^2) + \gamma (\delta(2; 1) + \delta(3; 3) + \delta(2; 3)) \\ &= 8,196 \end{aligned}$$

$$\begin{aligned} d(x_1, c_2) &= ((x_{11} - c_{21})^2 + (x_{12} - c_{22})^2) + \gamma (\delta(x_{13}, c_{23}) + \delta(x_{14}, c_{24}) + \delta(x_{15}, c_{25})) \\ &= ((15 - 12)^2 + (2 - 1)^2) + \gamma (\delta(2; 2) + \delta(3; 2) + \delta(2; 1)) \\ &= 14,196 \end{aligned}$$

$$\begin{aligned} d(x_1, c_3) &= ((x_{11} - c_{31})^2 + (x_{12} - c_{32})^2) + \gamma (\delta(x_{13}, c_{33}) + \delta(x_{14}, c_{34}) + \delta(x_{15}, c_{35})) \\ &= ((15 - 15)^2 + (2 - 3)^2) + \gamma (\delta(2; 3) + \delta(3; 1) + \delta(2; 2)) \\ &= 5,196 \end{aligned}$$

dan seterusnya sampai objek ke-65 $d(x_{65}, c_1)$, $d(x_{65}, c_2)$ dan $d(x_{65}, c_3)$.

Tabel 4.5 : Iterasi 1 Pengelompokan Hasil untuk Perhitungan Jarak dengan $k = 3$

Objek ke-i	$d(x_i, c_1)$	$d(x_i, c_2)$	$d(x_i, c_3)$	Nilai <i>Min</i>	<i>Cluster</i>
1	8,196	14,196	5,196	5,196	3
2	21,196	27,098	14,294	14,294	3
3	4,196	6,196	9,196	4,196	1
4	8,196	8,294	19,098	8,196	1
5	6,196	8,196	9,196	6,196	1
6	4,196	6,196	9,196	4,196	1
7	14,196	22,098	7,294	7,294	3
8	7,294	7,098	6,196	6,196	3
9	21,196	31,294	10,098	10,098	3
10	1	11,294	8,294	1	1

dan seterusnya sampai dengan 65 responden.

3. Dengan menggunakan persamaan 2.11 untuk variabel numerik dan 2.12 untuk variabel kategorik, seseorang dapat menentukan nilai *centroid* baru dari suatu objek yang termasuk dalam suatu kluster. 14 objek membentuk kluster 1, 5 objek membentuk kluster 2, dan 46 objek membentuk kluster 3 pada tahap kedua di atas. Nilai *centroid* baru dihitung sebagai berikut.

$$c_{11} = \frac{3+11+14+13+14+12+13+14+12+14+13+13+13+14}{14}$$

$$= 13,071$$

$$c_{12} = \frac{2+2+1+2+2+1+3+2+3+1+1+1+2+1}{14}$$

$$= 1,714$$

$$c_{13} = \text{Modus}(\{1; 3; \dots; 5\}) = 1$$

$$c_{14} = \text{Modus}(\{1; 3; \dots; 3\}) = 3$$

$$c_{15} = \text{Modus}(\{1; 2; \dots; 1\}) = 2$$

dan seterusnya sampai *cluster* 3.

Tabel 4.6 : Iterasi 2 dari Pengelompokan *Centroid* dengan $k = 3$

<i>Centroid</i>	Indeks Variabel ke- (p)				
	1	2	3	4	5
c_1	13,071	1,714	1	3	2
c_2	12,2	1	2	2	1
c_3	18,304	1,848	3	3	2

4. Hingga objek tetap diam dan nilai *centroid* dari dua iterasi sebelumnya tetap sama, prosedur dalam fase ini akan terus berulang. Iterasi ke-9 dari teknik ini berakhir dengan nilai *cluster* dan *centroid* yang sama.

Tabel 4.7 : Iterasi 9 dari Pengelompokan *Centroid* dengan $k = 3$

<i>Centroid</i>	Indeks Variabel ke- (p)				
	1	2	3	4	5
c_1	12,789	1,895	3	3	2
c_2	15,810	1,381	2	3	1
c_3	20,44	1,96	1	1	2

Iterasi 9

- Misalnya, rumus berikut dapat digunakan untuk menentukan jarak antara objek pertama dan masing-masing *centroid* kluster pertama (c_1), *centroid* kluster kedua (c_2), dan *centroid* kluster ketiga (c_3).

$$\begin{aligned}
d(x_1, c_1) &= ((x_{11} - c_{11})^2 + (x_{12} - c_{12})^2) + \gamma (\delta(x_{13}; c_{13}) + \delta(x_{14}; c_{14}) + \delta(x_{15}; c_{15})) \\
&= ((15 - 12, 789)^2 + (2 - 1, 895)^2) + \gamma (\delta(2; 3) + \delta(3; 3) + \delta(2; 3)) \\
&= 6,998
\end{aligned}$$

$$\begin{aligned}
d(x_1, c_2) &= ((x_{11} - c_{21})^2 + (x_{12} - c_{22})^2) + \gamma (\delta(x_{13}; c_{23}) + \delta(x_{14}; c_{24}) + \delta(x_{15}; c_{25})) \\
&= ((15 - 15, 810)^2 + (2 - 1, 381)^2) + \gamma (\delta(2; 2) + \delta(3; 3) + \delta(2; 1)) \\
&= 3,137
\end{aligned}$$

$$\begin{aligned}
d(x_1, c_3) &= ((x_{11} - c_{31})^2 + (x_{12} - c_{32})^2) + \gamma (\delta(x_{13}; c_{33}) + \delta(x_{14}; c_{34}) + \delta(x_{15}; c_{35})) \\
&= ((15 - 20, 44)^2 + (2 - 1, 96)^2) + \gamma (\delta(2; 1) + \delta(3; 1) + \delta(2; 2)) \\
&= 33,791
\end{aligned}$$

dan seterusnya sampai objek ke-65 $d(x_{65}, c_1)$, $d(x_{65}, c_2)$ dan $d(x_{65}, c_3)$.

Tabel 4.8 : Iterasi 1 Pengelompokan Hasil untuk Perhitungan Jarak dengan $k = 3$

Objek ke-i	$d(x_i, c_1)$	$d(x_i, c_2)$	$d(x_i, c_3)$	Nilai Min	Cluster
1	6,998	3,137	33,791	3,137	2
2	22,730	1,561	19,049	1,561	2
3	6,350	12,475	57,453	6,350	1
4	3,212	27,715	93,311	3,212	1
5	8,562	7,617	46,591	7,617	2
6	2,154	10,377	59,551	2,154	1
7	17,826	6,853	24,991	6,853	2
8	5,674	7,855	47,769	5,674	1
9	18,534	5,757	16,951	5,757	2
10	5,674	7,855	45,671	5,674	1

dan seterusnya sampai objek ke-65 $d(x_{65}, c_1)$, $d(x_{65}, c_2)$ dan $d(x_{65}, c_3)$.

Karakteristik masing-masing klaster ditunjukkan pada tabel 4.8, sedangkan tabel 4.7 menunjukkan nilai *centroid* pada iterasi ke-9:

a. *Cluster 1*

Cluster 1 dengan jumlah mahasiswa merokok yang terendah yaitu 19 mahasiswa. Prodi yang mahasiswanya kecanduan merokok yaitu matematika (x_3) dengan jenis rokok gudang garam (x_4). Jumlah rokok (x_2) yang dikonsumsi sebesar 2 bungkus dan awal usia merokok (x_1) mahasiswa rata-rata umur 12,789 tahun. Serta faktor yang menyebabkan merokok yaitu teman (x_5).

b. *Cluster 2*

Cluster 2 dengan jumlah mahasiswa merokok terbanyak dibanding *cluster 1* terdiri dari 21 mahasiswa. Prodi yang mahasiswanya kecanduan merokok

yaitu sistem informasi (x_3) dengan jenis rokok gudang garam (x_4). Jumlah rokok (x_2) yang dikonsumsi sebesar 1 bungkus dan awal usia merokok (x_1) mahasiswa rata-rata umur 15,810 tahun. Serta faktor yang menyebabkan merokok yaitu keluarga (x_5).

c. *Cluster 3*

Cluster 3 dengan jumlah mahasiswa merokok yang tertinggi yaitu 25 Mahasiswa. Prodi yang mahasiswanya kecanduan merokok yaitu ilmu komputer (x_3) dengan jenis rokok malboro (x_4). Jumlah rokok (x_2) yang dikonsumsi sebesar 2 bungkus dan awal usia merokok (x_1) mahasiswa rata-rata umur 20,44 tahun. Serta faktor yang menyebabkan merokok yaitu teman (x_5).

4.4 Pembahasan

Penelitian ini terdiri dari lima variabel yaitu variabel numerik (usia awal merokok dan jumlah rokok (bungkus)) dan variabel kategorik (prodi, jenis rokok dan faktor merokok). Pelaku kebiasaan merokok terjadi pada mahasiswa FST UINSU Medan, responden dalam penelitian ini adalah mahasiswa laki-laki tahun 2024. Mahasiswa merokok disebabkan oleh tiga faktor, yaitu keluarga, teman dan media massa (iklan rokok). Semua kelompok terutama remaja, dapat terpengaruh oleh maraknya iklan rokok di media. Salah satu faktor yang menyebabkan pola asuh orang tua yang buruk adalah maraknya remaja yang merokok. Remaja cenderung memiliki teman yang merokok jika mereka juga merokok. Fakta ini meningkatkan kemungkinan bahwa teman memiliki dampak terhadap remaja (susilaningsih, 2022).

Hasil penelitian ini memberikan informasi tentang pengelompokan data mahasiswa yang paling banyak merokok di FST UINSU Medan yang dapat diselesaikan dengan metode *K-Prototypes* setelah melalui proses penentuan variabel permasalahan, pengumpulan data, dan algoritma. Hasil yang didapat adalah tiga fakultas yang paling banyak kecanduan rokok yaitu ilmu komputer, sistem informasi, dan matematika. Perilaku merokok sebagian besar disebabkan oleh faktor seperti teman dan keluarga. Hasil penelitian terdahulu menunjukkan bahwa pengaruh pertemanan tidak relevan terhadap alasan perilaku merokok pada mahasiswa UINSU IV Tuntungan. Namun, berdasarkan hasil penelitian, perilaku merokok lebih sering dilakukan saat nongkrong atau bermain dengan teman.