

Social Engineering Vulnerability Index pada Provinsi di Indonesia Tahun 2025 dengan Principal Component Analysis dan K-Medoids Clustering

Hafizh Nursalim¹, Ratna Widayati^{2*}, Zulfia Memi Mayasari³, Oon Septa⁴

¹²³⁴Matematika, FMIPA, Universitas Bengkulu, Bengkulu

rwidayati@unib.ac.id

Diterima: 31-08-2026 ; Direvisi: 24-09-2026; Dipublikasi: 28-09-2026

Abstract

Indonesia generates a wide range of data, including both public data and sensitive personal data. Generally, cyber attacks can occur through software vulnerabilities, human intermediaries using social engineering techniques, or combination of both. If social engineering involves software vulnerabilities, CVSS can be used to assess software vulnerabilities and EPSS to estimate the probability of the exploitation. However, CVSS and EPSS cannot be applied to social engineering techniques that do not involve software or merely use it as a medium without exploiting its vulnerabilities. Based on literature review, no study has been found that develops Social Engineering Vulnerability Index at subnational level. Therefore, this research aims to fill this gap by constructing Social Engineering Vulnerability Index for 38 Indonesian provinces in 2025 using Principal Component Analysis (PCA), preserving 79.33% of original data's information. Subsequently, the vulnerability levels were analyzed using K-Medoids clustering, where the optimal number of clusters was based on interpretation of Silhouette Coefficient (SC) and average Silhouette Coefficient. The index interval is [0,1], with 5 vulnerability levels. Sensitivity analysis showed robust results regarding index ranking stability and varied cluster stability. This research indicates that there are variations in social engineering vulnerability due to socioeconomic and digital ecosystem disparities.

Keywords: Social Engineering Vulnerability; Index; Principal Component Analysis; K-Medoids Clustering

Abstrak

Indonesia menghasilkan beragam data, baik data publik maupun data pribadi yang bersifat sensitif. Secara umum, serangan cyber dapat terjadi melalui kerentanan perangkat lunak, melalui perantara sumber daya manusia dengan teknik *social engineering*, atau kombinasi keduanya. Jika *social engineering* melibatkan kerentanan perangkat lunak, maka dapat dilakukan penilaian CVSS untuk menilai kerentanan perangkat lunak dan EPSS untuk menilai probabilitas kerentanan tersebut dieksploitasi. Namun, CVSS dan EPSS tidak dapat diterapkan pada teknik *social engineering* yang tidak menggunakan perangkat lunak maupun menggunakan perangkat lunak sebagai sarana tanpa memanfaatkan kerentanannya. Berdasarkan kajian literatur, belum ditemukan penelitian yang membangun *Social Engineering Vulnerability Index* pada tingkat subnasional. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan membangun *Social Engineering Vulnerability Index* pada 38 provinsi di Indonesia tahun 2025 melalui *Principal Component Analysis* (PCA), yang mempertahankan 79,33% informasi data asli. Lalu, tingkat kerentanannya dianalisis dengan *K-Medoids clustering*, di mana jumlah *cluster* optimal didasarkan interpretasi *Silhouette Coefficient* (SC) dan rata-rata *Silhouette Coefficient*. Interval indeks adalah [0,1], dengan 5 tingkat kerentanan. Analisis sensitivitas menunjukkan hasil yang robust terhadap stabilitas peringkat indeks dan bervariasi terhadap stabilitas *cluster*. Penelitian ini menunjukkan bahwa terdapat variasi kerentanan terhadap *social engineering* karena kesenjangan sosial-ekonomi dan ekosistem digital.

Kata Kunci: Kerentanan *Social Engineering*; Indeks; *Principal Component Analysis*; *K-Medoids Clustering*

1. PENDAHULUAN

Indonesia adalah salah satu negara kepulauan terbesar di dunia yang memiliki 17.380 pulau (BIG, 2024) dan lebih dari 284,4 juta penduduk (BPS, 2025b). Karakteristik tersebut menjadikan Indonesia kaya akan keragaman budaya, etnis, dan kondisi geografis yang luar biasa. Setiap hari, berbagai aktivitas masyarakat Indonesia menghasilkan beragam data publik maupun data pribadi yang bersifat sensitif. Wawasan yang dihasilkan oleh data dapat berguna untuk mengetahui akar masalah dalam masyarakat, menentukan kebijakan yang sesuai, serta mengukur efektivitas program pembangunan secara akurat (Hossin et al., 2023). Namun keamanan data dan seluruh wawasannya, baik data publik maupun data pribadi yang bersifat sensitif, bergantung pada persiapan sistem keamanan yang memadai serta kualitas sumber daya manusia yang memiliki *security awareness* yang tinggi (Aldawood & Skinner, 2019).

Secara umum, ancaman terhadap keamanan data dan wawasannya dapat terjadi melalui perangkat lunak, melalui sumber daya manusia dengan teknik *social engineering*, atau kombinasi keduanya (Maraj & Butler, 2022). *Social engineering* merupakan teknik serangan *cyber* yang mengeksploitasi kerentanan psikologis dan kebiasaan manusia untuk mendapatkan akses tidak sah ke sistem atau informasi data (Kumar et al., 2024). Berdasarkan skema terjadinya ancaman keamanan data dan wawasannya, teknik *social engineering* dapat terjadi melalui interaksi langsung tanpa melibatkan perangkat lunak atau melalui interaksi dengan melibatkan perangkat lunak. Meskipun memiliki skema yang berbeda, keduanya sama-sama bertumpu pada eksploitasi manusia sebagai titik lemah utama. Hal ini tercermin dalam *Data Breach Investigation Report* (DBIR) yang menunjukkan bahwa 82% dari serangan *cyber* yang tercatat melibatkan faktor manusia sebagai pusat perhatian dalam serangan (Bassett et al., 2022).

Pada skema serangan *social engineering* yang melibatkan perangkat lunak, keterlibatan perangkat lunak dapat terjadi dalam dua kondisi. Kondisi pertama terjadi dengan melibatkan kerentanan perangkat lunak, sedangkan kondisi kedua hanya menggunakan perangkat lunak sebagai sarana tanpa memanfaatkan kerentanan sama sekali. Jika teknik *social engineering* melibatkan kerentanan perangkat lunak, maka dapat dilakukan penilaian *Common Vulnerability Scoring System* (CVSS). Namun, CVSS yang dikelola oleh *Forum of Incident Response and Security Teams* (FIRST), hanya mampu mengukur tingkat keparahan dari kerentanan perangkat lunak dan bukan termasuk ukuran probabilitas (FIRST, 2023). Untuk melengkapi celah CVSS tersebut, Jacobs et al. (2021) telah merancang *Exploit Prediction Scoring System* (EPSS) untuk mengukur probabilitas *Common Vulnerabilities and Exposures* (CVE) akan dieksploitasi dari suatu perangkat lunak. Kemudian, EPSS dikembangkan dan dikelola lebih lanjut oleh FIRST. Namun, baik CVSS maupun EPSS hanya dirancang untuk kerentanan perangkat lunak. Oleh karena itu, CVSS dan EPSS tidak dapat diterapkan pada teknik *social engineering* yang tidak melibatkan perangkat lunak maupun menggunakan perangkat lunak sebagai sarana tanpa memanfaatkan kerentanan sama sekali.

Berbeda dengan kerangka kerja CVSS dan EPSS dengan fokus kerentanan perangkat lunak, sejumlah penelitian telah berupaya untuk mengkaji dan mengukur kerentanan terhadap *social engineering*. Wang et al. (2012) menemukan bahwa kerentanan individu terhadap serangan *spear phishing* meningkat akibat adanya isyarat instingtif, sementara pengetahuan penipuan digital menjadi faktor pelindung yang signifikan. Vishwanath et al. (2016) berhasil membangun *Suspicion, Cognition, and Automaticity Model* (SCAM) yang menunjukkan bahwa kebiasaan individu dalam menggunakan email secara refleks tanpa pertimbangan kognitif membuat individu tidak sempat merasa curiga, sehingga lebih rentan terhadap *phishing*. Namun, Wang et al. (2012) dan Vishwanath et al. (2016) hanya meneliti sebagian skenario teknik *social engineering* dan menggunakan sampel mahasiswa sehingga sulit mewakili masyarakat umum dengan beragam latar belakang sosial-ekonomi. Kemudian, Albladi & Weir (2020) memprediksi kerentanan pengguna media sosial terhadap *social engineering* berdasarkan kepercayaan sebagai prediktor terkuat, diikuti tingkat keterlibatan, motivasi, dan kompetensi. Namun, Albladi & Weir (2020) menggunakan eksperimen berbasis skenario buatan dengan sampel komunitas akademik dalam universitas. Dalam bidang pemodelan matematika, Şandor et al. (2022) mengusulkan kerangka penilaian risiko organisasi terhadap *social engineering* berbasis simulasi Monte Carlo. Namun, kerangka penilaian dari Şandor et al. (2022) memiliki fleksibilitas terbatas karena sangat bergantung pada estimasi probabilitas satu skenario yang spesifik. Lebih lanjut, Mitra et al. (2025) membangun *Social Cyber Vulnerability Index* (SCVI) melalui integrasi faktor individu dan karakteristik serangan. Meskipun demikian, Mitra et al. (2025) menghadapi tantangan karena menggunakan data survei melalui iPoll dan Reddit sehingga hanya dapat diakses oleh kelompok orang tertentu, serta dapat terjadi risiko bias ingatan dan kurangnya pelaporan dalam survei.

Berdasarkan kajian literatur yang telah dilakukan, belum ditemukan penelitian yang membangun indeks kerentanan masyarakat terhadap *social engineering* pada tingkat subnasional. Oleh karena itu, penelitian ini bertujuan untuk membangun *Social Engineering Vulnerability Index* dengan memanfaatkan variabel sosial-ekonomi dan teknologi yang tersedia secara terbuka. Pembangunan indeks dilakukan menggunakan *Principal Component Analysis* (PCA) dengan mempertahankan sebagian besar informasi yang terdapat pada data asli. Selanjutnya, tingkat kerentanan dianalisis menggunakan *K-Medoids clustering* dengan jumlah *cluster* optimal ditentukan berdasarkan interpretasi *Silhouette Coefficient* dan rata-ratanya. Hasil penelitian ini diharapkan dapat menjadi dasar dalam mengidentifikasi tingkat kerentanan provinsi di Indonesia terhadap *social engineering* sehingga dapat mendukung penyusunan kebijakan, peningkatan literasi keamanan *cyber*, dan penentuan prioritas program mitigasi secara lebih akurat.

2. METODE PELAKSANAAN

Penelitian ini dilakukan di Universitas Bengkulu dengan pendekatan kuantitatif. Subjek penelitian ini adalah 38 provinsi di Indonesia pada tahun 2025 sehingga konsistensi dan kebaruan data terjaga. Penelitian ini dilakukan dengan tahapan utama sebagai berikut:

2.1 Proses Pengumpulan Data

Pada tahapan pertama, proses pengumpulan data diambil dari sumber yang bersifat kredibel dan tersedia secara publik. Data penelitian mencakup 6 variabel di mana setiap variabel berbentuk vektor kolom dengan 38 baris sebagai representasi nilai pada setiap provinsi. Dalam pengumpulan data, variabel yang dikumpulkan meliputi:

1. Tingkat penetrasi internet ($\overline{X_1}$) (APJII, 2025), sebagai ukuran paparan *digital* terhadap ancaman *cyber*.
2. Rata-rata lama sekolah penduduk berusia di atas 15 tahun ($\overline{X_2}$) (BPS, 2025e), yang berkaitan dengan kemampuan literasi dan pemahaman informasi masyarakat.
3. Persentase penduduk miskin ($\overline{X_3}$) (BPS, 2025d), yang berkaitan dengan kerentanan ekonomi dan psikologis masyarakat.
4. *Gini ratio* ($\overline{X_4}$) (BPS, 2025a), sebagai ukuran ketimpangan distribusi ekonomi yang memiliki potensi ketidakpuasan sosial.
5. Tingkat Pengangguran Terbuka (TPT) ($\overline{X_5}$) (BPS, 2025f), yang berkaitan dengan kerentanan masyarakat terhadap *social engineering* dengan modus pekerjaan.
6. Kepadatan penduduk ($\overline{X_6}$) (BPS, 2025c), sebagai ukuran kecepatan penyebaran informasi dalam masyarakat.

Setelah data diperoleh, data disimpan dalam bentuk tabel pada *file* Excel, di mana kolom pertama mencakup 38 nama provinsi di Indonesia, serta kolom kedua hingga ketujuh mencakup 38 nilai dari variabel $\overline{X_1}$ hingga $\overline{X_6}$. Kemudian, *file* Excel di-*import* pada program RStudio versi 4.4.2 dan diproses untuk pengolahan data pada tahap berikutnya.

2.2 Principal Component Analysis (PCA)

Pada tahapan kedua, data diolah menggunakan *Principal Component Analysis* (PCA) untuk membangun *Social Engineering Vulnerability Index* sehingga mampu mereduksi sejumlah variabel yang saling berkorelasi dengan tetap mempertahankan sebagian besar informasi pada data asli. Dalam proses membangun *Social Engineering Vulnerability Index* pada 38 provinsi di Indonesia tahun 2025, rincian tahapan yang dilakukan adalah sebagai berikut:

- 1 Melakukan *preprocessing data* dengan menjadikan kolom pertama sebagai *rownames*, sehingga hanya variabel numerik $\overline{X_1}$ hingga $\overline{X_6}$ yang diproses dalam perhitungan. Melalui variabel $\overline{X_a}$ untuk $a = 1, 2, 3, \dots, 6$, dapat dibentuk matriks $\mathbf{X}$ yang memiliki elemen $x_{b,a}$ untuk $b = 1, 2, 3, \dots, 38$. Elemen $x_{b,a}$ dianggap sebagai nilai variabel $\overline{X_a}$ pada provinsi ke- b . Matriks $\mathbf{X}$ yang dibentuk terdapat pada Persamaan (1) sebagai berikut:

$$\mathbf{X} = \begin{bmatrix} x_{1.1} & x_{1.2} & x_{1.3} & \dots & x_{1.6} \\ x_{2.1} & x_{2.2} & x_{2.3} & \dots & x_{2.6} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{38.1} & x_{38.2} & x_{38.3} & \dots & x_{38.6} \end{bmatrix} \quad (1)$$

- 2 Variabel $\overline{X}_a$ dilakukan standardisasi Z menjadi variabel $\overline{Z}_a$ untuk $a = 1, 2, 3, \dots, 6$ sehingga elemen $x_{b,a}$ distandardisasi menjadi elemen $z_{b,a}$. Elemen $z_{b,a}$ dianggap sebagai nilai variabel $\overline{Z}_a$ pada provinsi ke- b untuk $b = 1, 2, 3, \dots, 38$.
- 3 Melalui elemen $z_{b,a}$, dapat dibentuk matriks $\mathbf{Z}$ seperti pada Persamaan (2) berikut:

$$\mathbf{Z} = \begin{bmatrix} z_{1.1} & z_{1.2} & z_{1.3} & \dots & z_{1.6} \\ z_{2.1} & z_{2.2} & z_{2.3} & \dots & z_{2.6} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{38.1} & z_{38.2} & z_{38.3} & \dots & z_{38.6} \end{bmatrix} \quad (2)$$

- 4 Membentuk matriks *covariance* (Σ_Z) dari variabel $\overline{Z}_a$ yang *equivalen* terhadap matriks korelasi dari variabel $\overline{X}_a$ untuk $a = 1, 2, 3, \dots, 6$.
- 5 Mencari nilai *eigen* (λ_a) dan vektor *eigen* ($\overline{v}_a$) dari matriks Σ_Z untuk $a = 1, 2, 3, \dots, 6$.
- 6 Untuk $a = 1, 2, 3, \dots, 6$, nilai *eigen* (λ_a) dan vektor *eigen* ($\overline{v}_a$) diurutkan berdasarkan nilai *eigen* (λ_a) dari terbesar hingga terkecil.
- 7 Untuk $a = 1, 2, 3, \dots, 6$, dihitung persentase kontribusi nilai *eigen* pada indeks ke- a (c_a) terhadap variansi data melalui Persamaan (3) sebagai berikut:

$$c_a = \frac{\lambda_a}{\sum_{i=1}^6 \lambda_i} * 100\% \quad (3)$$

- 8 Untuk $a = 1, 2, 3, \dots, 6$, dipilih n nilai *eigen* (λ_a) terbesar yang memiliki kontribusi besar melalui kriteria Kaiser Guttman untuk mempertahankan $\lambda_a > 1$ (Kaiser, 1960), dan dikombinasikan dengan visualisasi *scree plot* untuk mempertimbangkan total persentase kontribusi nilai *eigen* (c_a) terhadap variansi data. Selanjutnya, vektor *eigen* ($\overline{v}_a$) juga dipilih berdasarkan nilai *eigen* (λ_a) tersebut.
- 9 Untuk $a = 1, 2, 3, \dots, n$, vektor *Principal Component* ($\overline{PC}_a$) diperoleh melalui:

$$\overline{PC}_a = \mathbf{Z} * \overline{v}_a \quad (4)$$

Kemudian dibentuk matriks *Principal Component* ($\mathbf{PC}$) yang berisikan setiap vektor *Principal Component* ($\overline{PC}_a$) secara berurutan pada setiap kolom matriks.

- 10 Membentuk vektor bobot ($\overline{w}$) berdasarkan proporsi kontribusi variansi yang berbeda dengan mengacu pada pendekatan yang digunakan oleh Trivianita et al. (2019) melalui Persamaan (5) berikut:

$$\overline{w} = \begin{bmatrix} \frac{\lambda_1}{\sum_{i=1}^n \lambda_i} \\ \frac{\lambda_2}{\sum_{i=1}^n \lambda_i} \\ \vdots \\ \frac{\lambda_n}{\sum_{i=1}^n \lambda_i} \end{bmatrix} \quad (5)$$

- 11 Mengonstruksi *Social Engineering Vulnerability Index* ($\overline{SEVI}$) dengan mengacu pada pendekatan yang digunakan oleh Trivianita et al. (2019) melalui Persamaan (6) sebagai berikut:

$$\overline{SEVI} = \mathbf{PC} * \overline{w} \quad (6)$$

Setiap elemen pada vektor $\overline{SEVI}$ berisi nilai indeks suatu provinsi di Indonesia dengan total 38 baris.

- 12 Menstandarisasi setiap nilai elemen $\overline{SEVI}$ dari 0 hingga 1 dengan menggunakan metode *min-max* sehingga diperoleh $\overline{SEVI}'$.

2.3 *K-Medoids Clustering*

Pada tahapan ketiga, metode *K-Medoids clustering* digunakan untuk mengetahui *cluster* kerentanan provinsi berdasarkan karakteristik setiap elemen $\overline{SEVI}'$ sebagai nilai indeks akhir. Metode ini dipilih karena lebih *robust* terhadap nilai *outlier*. Pada tahapan awal *K-Medoids clustering*, penentuan jumlah *cluster* optimal didasarkan pada pertimbangan interpretasi dari *Silhouette Coefficient* (*SC*) dan rata-rata *Silhouette Coefficient*. Nilai *Silhouette Coefficient* (*SC*) digunakan untuk melihat seberapa baik suatu provinsi berada dalam *cluster*-nya dibandingkan dengan *cluster* lain terdekat. Semakin besar nilai dari *Silhouette Coefficient* (*SC*) yang berkisar antara -1 hingga 1 , semakin baik suatu provinsi berada dalam *cluster*-nya dibandingkan *cluster* lain terdekat dengan interpretasi berikut:

- 1 Jika $SC > 0$, maka posisi provinsi pada suatu *cluster* lebih baik dibandingkan dengan *cluster* lain terdekat
- 2 Jika $SC = 0$, maka posisi provinsi pada suatu *cluster* bersifat ambigu
- 3 Jika $SC < 0$, maka posisi provinsi pada suatu *cluster* kurang baik dibandingkan dengan *cluster* lain terdekat

Selanjutnya, nilai rata-rata *Silhouette Coefficient* digunakan untuk melihat kekuatan struktur *cluster* pada beberapa nilai K . Berdasarkan Kaufman & Rousseeuw (1990), interpretasi rata-rata *Silhouette Coefficient* terdapat pada Tabel 1 sebagai berikut:

Tabel 1. Interpretasi Rata-Rata *Silhouette Coefficient*

Rata-Rata <i>Silhouette Coefficient</i>	Interpretasi Struktur Cluster
0,71 – 1,00	Struktur kuat
0,51 – 0,70	Struktur baik
0,26 – 0,50	Struktur lemah
$\leq 0,25$	Tidak ada struktur

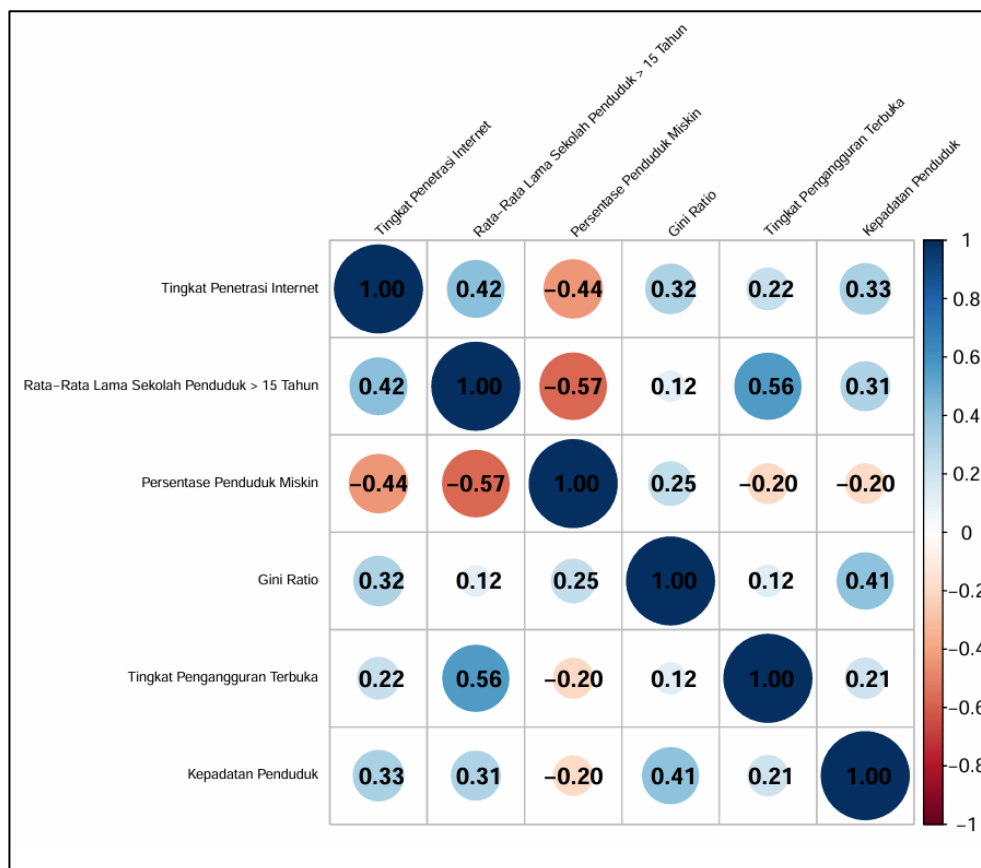
Setelah nilai *Silhouette Coefficient* (*SC*) dan rata-rata *Silhouette Coefficient* diperoleh, penentuan jumlah *cluster* (K) optimal dilakukan dengan membandingkan nilai rata-rata *Silhouette Coefficient* pada beberapa skenario nilai K . Nilai K yang dipilih adalah nilai yang memiliki rata-rata *Silhouette Coefficient* tertinggi tanpa adanya provinsi dengan *Silhouette Coefficient* (*SC*) negatif. Setelah nilai K optimal diperoleh, algoritma *K-Medoids clustering* diterapkan terhadap *Social Engineering Vulnerability Index* sehingga setiap provinsi memiliki *cluster*-nya masing-masing yang mewakili tingkatan kerentanan *social engineering*. Pada hasil program RStudio, label angka yang tertera pada setiap *cluster* belum tentu merepresentasikan urutan tingkat kerentanan secara langsung. Mengingat proses *clustering* hanya dilakukan terhadap satu variabel tunggal, maka tidak akan ada tumpang tindih rentang nilai antar *cluster*. Oleh karena itu, penentuan urutan tingkat

kerentanan direpresentasikan dengan posisi rentang nilai indeks suatu *cluster* terhadap *cluster* lainnya. Sebaran indeks dengan informasi tingkatan kerentanan dan *outlier* dari indeks direpresentasikan oleh *boxplot*. Berdasarkan hasil penelitian, analisis sensitivitas dilakukan terhadap jumlah *principal component* dan jumlah *cluster* pada Subbab 3.7.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Korelasi dari Variabel Penelitian

Setelah melalui tahapan 1 hingga 3 pada metode *Principal Component Analysis* (PCA), dapat dilakukan tahapan 4 sehingga diperoleh matriks *covariance* (Σ_Z) dari variabel $\bar{Z}_a$ untuk $a = 1, 2, 3, \dots, 6$. Matriks *covariance* (Σ_Z) dari variabel $\bar{Z}_a$ *equivalen* terhadap matriks korelasi dari variabel $\bar{X}_a$ untuk $a = 1, 2, 3, \dots, 6$. Berdasarkan matriks *covariance* (Σ_Z) dari variabel $\bar{Z}_a$ untuk $a = 1, 2, 3, \dots, 6$, dapat direpresentasikan korelasi dari setiap variabel penelitian seperti pada Gambar 1 sebagai berikut:



Gambar 1. Visualisasi Korelasi dari Variabel Penelitian

3.2 Hasil Nilai *Eigen* dan Vektor *Eigen* dari Matriks *Covariance* (Σ_Z)

Setelah melalui tahapan 4 pada *Principal Component Analysis* (PCA), dapat dilakukan tahapan 5 hingga 6 sehingga diperoleh nilai *eigen* (λ_a) dan vektor *eigen* ($\bar{v}_a$) dari matriks

covariance (Σ_Z) berdasarkan urutan nilai *eigen* (λ_a) secara menurun untuk $a = 1, 2, 3, \dots, 6$ sebagai berikut:

Tabel 2. Nilai *Eigen* dari Matriks Covariance (Σ_Z)

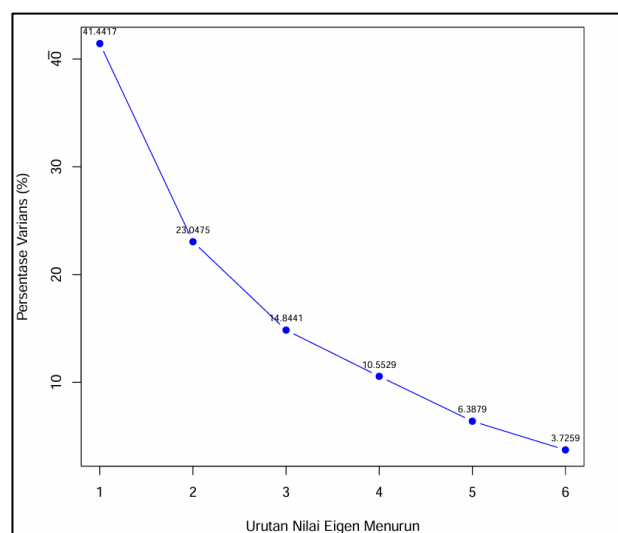
λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
2,486502018	1,382850079	0,890647522	0,633171937	0,383271686	0,223556758

Tabel 3. Vektor *Eigen* dari Matriks Covariance (Σ_Z)

$\bar{v}_1$	$\bar{v}_2$	$\bar{v}_3$	$\bar{v}_4$	$\bar{v}_5$	$\bar{v}_6$
-0,4563421	-0,09569118	0,44739935	-0,56279941	0,44270251	-0,26400787
-0,53046121	0,20861945	-0,22455702	-0,00115749	-0,58991675	-0,52598464
0,4095807	-0,50266102	-0,37470369	-0,1226952	0,1597391	-0,63134927
-0,20761140	-0,73178190	-0,03482650	-0,26305792	-0,42569768	0,41202168
-0,39619941	0,06911186	-0,75789255	-0,07337638	0,43555571	0,26221282
-0,3775312	-0,39290177	0,18273228	0,77046595	0,24406721	-0,12853270

3.3 Hasil Pemilihan Nilai *Eigen* dan Vektor *Eigen*

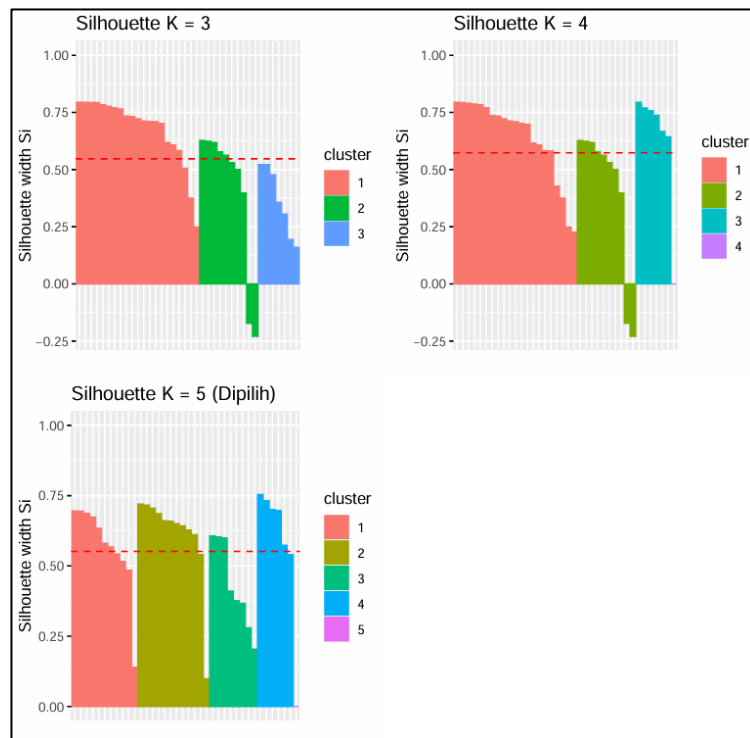
Setelah melalui tahapan 6 pada metode *Principal Component Analysis* (PCA), dapat dilakukan tahapan 7 hingga 8 untuk memilih n nilai *eigen* (λ_a) terbesar dan vektor *eigen* ($\bar{v}_a$) yang sesuai dengan hasil $n = 3$. Pemilihan tersebut didasarkan pada kriteria Kaiser-Guttman (Kaiser, 1960) dan dikombinasikan dengan visualisasi *scree plot* dari persentase kontribusi nilai *eigen* pada indeks ke- a (c_a) terhadap variansi data. Adapun 3 nilai *eigen* terbesar yang dipilih yaitu λ_a untuk $a = 1, 2, 3$ pada Tabel 2, sedangkan vektor *eigen* yang dipilih yaitu $\bar{v}_a$ untuk $a = 1, 2, 3$ pada Tabel 3. Nilai *eigen* (λ_a) untuk $a = 1, 2, 3$ secara kumulatif menjelaskan 79,33% total variansi data. Visualisasi *scree plot* dari kontribusi nilai *eigen* (c_a) terhadap variansi data untuk $a = 1, 2, 3, \dots, 6$ adalah sebagai berikut:



Gambar 2. Visualisasi Kontribusi Nilai *Eigen* Terhadap Variansi Data

3.4 Hasil Pemilihan Jumlah *Cluster* Optimal

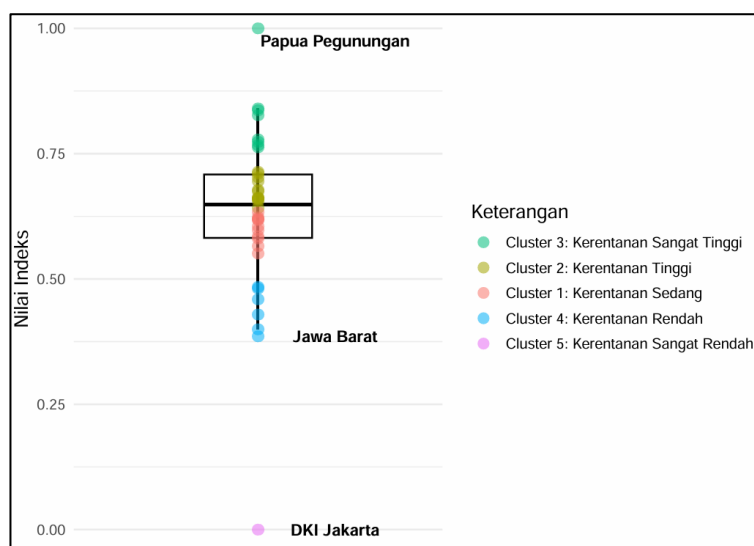
Proses pemilihan jumlah *cluster* optimal dilakukan setelah tahapan *Principal Component Analysis* (PCA) menghasilkan $\overline{SEVI}$ dengan detail nilai yang tertera pada Tabel 4 di Subbagian 3.5. Pada pemilihan jumlah *cluster* optimal, diperlukan nilai *Silhouette Coefficient* (SC) dan rata-rata *Silhouette Coefficient*. Berdasarkan proses tersebut, dipilih jumlah *cluster* $K = 5$. Visualisasi *Silhouette Coefficient* (SC) dan rata-rata *Silhouette Coefficient* untuk beberapa skenario nilai K terdapat pada Gambar 3 sebagai berikut:



Gambar 3. Visualisasi *Silhouette Coefficient* dan Rata-Rata *Silhouette Coefficient*

3.5 Hasil Cluster Indeks

Setelah nilai $K = 5$ diterapkan pada proses *K-Medoids clustering*, dapat diperoleh *cluster* yang mewakili tingkat kerentanan *social engineering* dari setiap 38 provinsi di Indonesia tahun 2025. Berdasarkan Subbagian 2.3, penentuan urutan tingkat kerentanan dapat direpresentasikan melalui posisi rentang indeks suatu *cluster* terhadap *cluster* lainnya. Berikut ini sebaran nilai indeks dengan informasi tingkatan kerentanan dan *outlier*-nya:



Gambar 4. Boxplot Indeks dan Tingkat Kerentanannya

Lebih lanjut, detail nilai indeks dan *cluster* tingkat kerentanan setiap 38 provinsi di Indonesia tahun 2025 terdapat pada Tabel 4 sebagai berikut:

Tabel 4. Nilai dan *Cluster* Tingkat Kerentanan dari Indeks

No	Provinsi	Indeks
Tingkat Kerentanan Sangat Tinggi (<i>Cluster 3</i>)		
1	Papua Pegunungan	1
2	Sulawesi Barat	0,840239849
3	Papua Tengah	0,836650183
4	Sulawesi Tengah	0,826765709
5	Kepulauan Bangka Belitung	0,778400522
6	Kalimantan Utara	0,773545593
7	Kalimantan Selatan	0,767025653
8	Nusa Tenggara Timur	0,76282747
Tingkat Kerentanan Tinggi (<i>Cluster 2</i>)		
9	Lampung	0,713675479
10	Sumatera Selatan	0,710370161
11	Maluku Utara	0,702978278
12	Nusa Tenggara Barat	0,695110237
13	Kalimantan Barat	0,677675963
14	Kalimantan Tengah	0,675071926
15	Bali	0,662622105
16	Riau	0,661407415
17	Gorontalo	0,660697426
18	Jambi	0,658658602
19	Sulawesi Tenggara	0,655161581

No	Provinsi	Indeks
20	Sumatera Utara	0,641665089
Tingkat Kerentanan Sedang (<i>Cluster 1</i>)		
21	Sumatera Barat	0,633371437
22	Sulawesi Selatan	0,622941601
23	Aceh	0,620295091
24	Bengkulu	0,618761603
25	Jawa Timur	0,61770981
26	Maluku	0,604935828
27	Jawa Tengah	0,597502388
28	Papua Selatan	0,585122444
29	Papua Barat	0,580853556
30	Kalimantan Timur	0,565611933
31	Sulawesi Utara	0,550927538
Tingkat Kerentanan Rendah (<i>Cluster 4</i>)		
32	Banten	0,484144905
33	Papua Barat Daya	0,481009497
34	DI Yogyakarta	0,45940498
35	Kepulauan Riau	0,429087018
36	Papua	0,399620456
37	Jawa Barat	0,385647053
Tingkat Kerentanan Sangat Rendah (<i>Cluster 5</i>)		
38	DKI Jakarta	0

3.6 Pembahasan

Pada Subbagian 3.1, hasil korelasi dari variabel penelitian diperlukan dalam tahapan *Principal Component Analysis* (PCA) untuk memastikan adanya hubungan antar variabel. Berdasarkan Gambar 1 yang merepresentasikan matriks korelasi, nilai korelasi dari variabel yang berbeda cenderung tidak bernilai 0 atau tidak mendekati 0. Oleh sebab itu, matriks korelasi dari variabel penelitian bukan termasuk matriks identitas. Kondisi ini menandakan bahwa adanya hubungan antar variabel, sehingga syarat dalam melakukan *Principal Component Analysis* (PCA) telah terpenuhi untuk menangkap pola hubungan yang ada. Semakin besar jarak antara nilai korelasi dari dua variabel terhadap 0 dengan batas maksimal 1, semakin kuat juga hubungan antara kedua variabel tersebut. Lebih lanjut, hubungan antara persentase penduduk miskin dan rata-rata lama sekolah penduduk di atas 15 tahun menunjukkan korelasi negatif terkuat dengan nilai $-0,57$ mengindikasikan semakin tinggi rata-rata lama sekolah penduduk, semakin rendah persentase penduduk miskin, begitu juga sebaliknya. Hal ini juga berlaku kebalikan untuk korelasi positif. Selain itu, variabel lain juga menunjukkan pola

hubungan yang bervariasi, menegaskan bahwa secara keseluruhan terdapat keterkaitan antar variabel penelitian dengan tingkat kekuatan yang berbeda-beda.

Pada Subbagian 3.3, pemilihan n nilai *eigen* (λ_a) terbesar dan vektor *eigen* ($\overline{v_a}$) yang sesuai didasarkan pada kriteria Kaiser-Guttman (Kaiser, 1960) dan visualisasi *scree plot* pada Gambar 2. Berdasarkan Tabel 2 di Subbagian 3.2, nilai *eigen* (λ_a) yang memenuhi kriteria Kaiser-Guttman yang mempertahankan $\lambda_a > 1$ adalah $\lambda_1 = 2,486502018$ dan $\lambda_2 = 1,382850079$. Berdasarkan visualisasi *scree plot* pada Gambar 2 di Subbagian 3.3, persentase kontribusi nilai *eigen* λ_1 dan λ_2 terhadap variansi data secara berurutan yaitu $c_1 = 41,44\%$ dan $c_2 = 23,05\%$. Secara kumulatif, pemilihan λ_1 dan λ_2 dapat menjelaskan 64,49% dari variansi data. Kemudian, persentase kontribusi nilai *eigen* $\lambda_3 = 0,890647522$ terhadap variansi data yaitu 14,84%. Secara kumulatif, pemilihan λ_1 , λ_2 , dan λ_3 dapat menjelaskan 79,33% dari variansi data. Perbedaan yang cukup besar antara nilai 79,33% terhadap 64,49% merepresentasikan informasi dari data yang dapat dipertahankan. Oleh karena itu, nilai *eigen* dan vektor *eigen* yang dipilih adalah λ_1 , λ_2 , λ_3 dan $\overline{v_1}$, $\overline{v_2}$, $\overline{v_3}$. Pemilihan ini dilakukan dengan tetap mempertahankan λ_1 dan λ_2 yang memenuhi kriteria Kaiser-Guttman. Capaian ini menunjukkan bahwa pemilihan yang dilakukan sudah cukup representatif tanpa kehilangan informasi secara berarti.

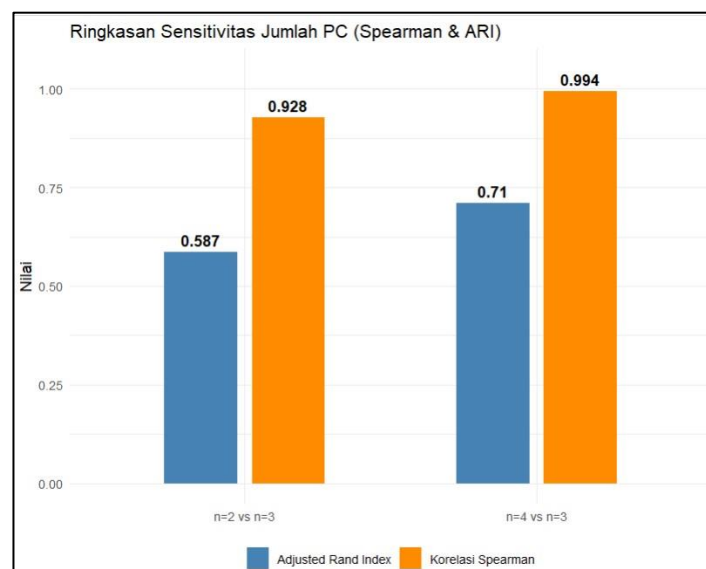
Pada Subbagian 3.4, pemilihan jumlah *cluster* optimal dilakukan pada beberapa skenario nilai K yang divisualisasikan pada Gambar 3. Skenario nilai $K > 5$ tidak dievaluasi karena menghasilkan tingkat kerentanan yang terlalu kompleks untuk tujuan kebijakan. Sementara itu, skenario nilai $K = 1$ dan $K = 2$ juga tidak dievaluasi karena tingkat kerentanan yang dihasilkan dinilai tidak dapat menggambarkan variasi tingkat kerentanan secara rinci. Oleh karena itu, digunakan perbandingan untuk memilih skenario nilai $K = 3$, $K = 4$, atau $K = 5$. Berdasarkan visualisasi pada Gambar 3, nilai rata-rata *Silhouette Coefficient* untuk $K = 3$ hingga $K = 5$ berada dalam rentang 0,51 hingga 0,70 sehingga memiliki struktur *cluster* yang baik untuk setiap nilai K yang diuji. Namun, nilai $K = 3$ dan $K = 4$ memiliki provinsi dengan *Silhouette Coefficient* (SC) yang bernilai negatif sehingga terdapat provinsi dengan *cluster* yang tidak sesuai. Sebaliknya, untuk nilai $K = 5$, seluruh provinsi tidak memiliki nilai *Silhouette Coefficient* (SC) negatif sehingga setiap provinsi memiliki *cluster* yang sesuai. Berdasarkan hal tersebut, dipilih skenario $K = 5$ yang merepresentasikan tingkatan kerentanan sangat tinggi, tinggi, sedang, rendah, dan sangat rendah.

Pada Subbagian 3.5, nilai indeks dan tingkatan kerentanan pada 38 provinsi di Indonesia terdapat pada Tabel 4. Untuk mengetahui sebaran nilai indeks, dapat dilihat posisi rentang suatu *cluster* terhadap *cluster* lainnya. Berdasarkan *boxplot* pada Gambar 4, posisi suatu *cluster* berdasarkan rentang nilai indeks terbesar hingga terkecil yaitu *cluster* 3, *cluster* 2, *cluster* 1, *cluster* 4, dan *cluster* 5 sehingga merepresentasikan tingkatan kerentanan sangat tinggi, tinggi, sedang, rendah, dan sangat rendah. Visualisasi *boxplot* pada Gambar 4 menunjukkan bahwa terdapat *outlier* data yaitu pada

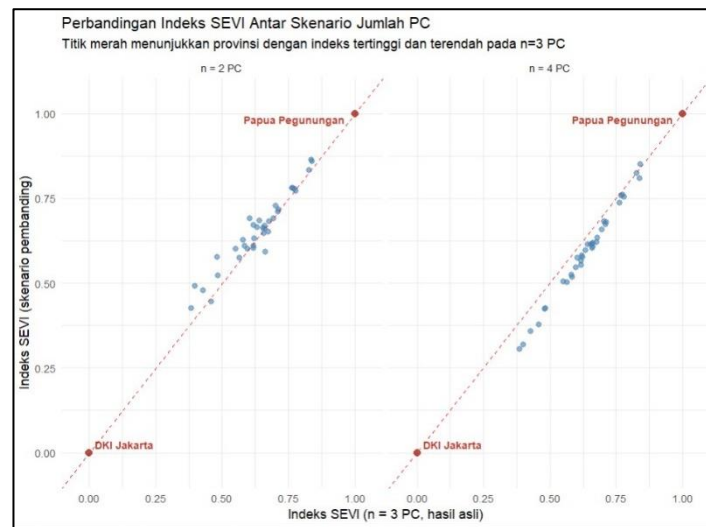
provinsi DKI Jakarta dengan tingkat kerentanan sangat rendah, Jawa Barat dengan tingkat kerentanan rendah, dan papua pegunungan dengan tingkat kerentanan sangat tinggi. Berdasarkan visualisasi *boxplot* pada Gambar 4 dengan detail pada Tabel 4, untuk provinsi dengan nilai indeks terendah, terdapat perbedaan nilai yang mencolok antara DKI Jakarta terhadap Jawa Barat sebagai provinsi dengan nilai indeks terdekat. Sementara itu, untuk provinsi yang memiliki nilai indeks tertinggi, terdapat perbedaan nilai yang mencolok antara Papua Pegunungan terhadap Sulawesi Barat sebagai provinsi yang memiliki nilai indeks terdekat. Hasil temuan tersebut dikarenakan kondisi sosial-ekonomi dan ekosistem *digital* pada setiap provinsi di Indonesia masih belum merata dalam menghadapi *social engineering*. Secara keseluruhan, penelitian ini telah menjawab tujuan yang telah dirumuskan yaitu untuk membangun *Social Engineering Vulnerability Index* pada tingkat subnasional yang selama ini belum banyak dikaji dalam literatur. Penelitian ini diharapkan dapat memberikan kontribusi untuk mendukung dalam penyusunan kebijakan secara akurat untuk mengatasi kesenjangan kondisi sosial ekonomi dan ekosistem *digital* dalam menghadapi serangan *social engineering*.

3.7 Analisis Sensitivitas

Analisis sensitivitas terhadap pemilihan jumlah *principal component* dilakukan dengan menghitung ulang indeks menggunakan skenario 2 *principal component* ($n = 2$) dan 4 *principal component* ($n = 4$). Jumlah *cluster* yang digunakan pada setiap skenario tetap bernilai $K = 5$. Hasil perhitungan indeks setiap skenario dibandingkan terhadap indeks awal yang hanya menggunakan 3 *principal component* ($n = 3$). Konsistensi peringkat provinsi dievaluasi berdasarkan hasil dari korelasi Spearman, sedangkan konsistensi kelompok *cluster* dievaluasi berdasarkan *Adjusted Rand Index* (ARI). Detail dari hasil analisis sensitivitas tersebut terdapat pada Gambar 5 dan Gambar 6 sebagai berikut:



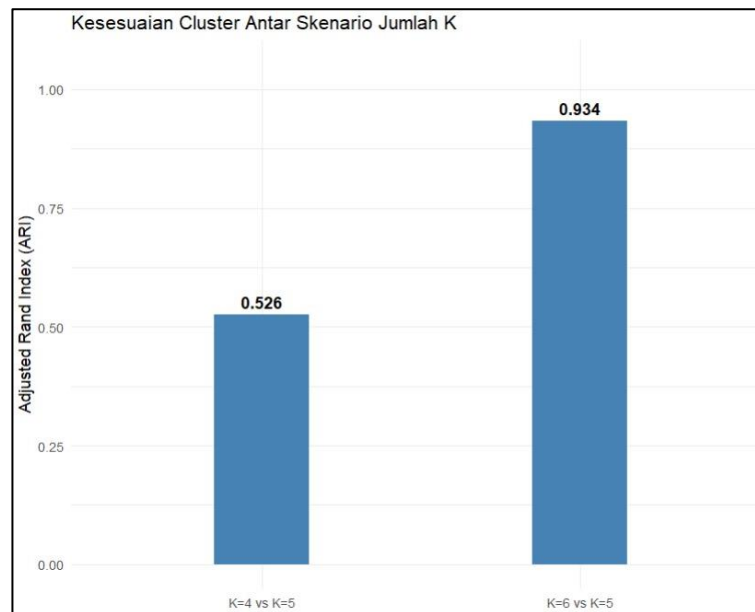
Gambar 5. Barchart Korelasi Spearman dan ARI pada Analisis Sensitivitas Pemilihan Jumlah *Principal Component*



Gambar 6. Scatter Plot Perbandingan Indeks Awal Terhadap Indeks Pembanding pada Analisis Sensitivitas Pemilihan Jumlah *Principal Component*

Berdasarkan Gambar 5, korelasi Spearman antara skenario $n = 2$ dan $n = 4$ terhadap $n = 3$ melampaui batas ambang 0,90. Hal tersebut mengindikasikan bahwa mayoritas peringkat indeks provinsi bersifat *robust* terhadap perubahan dari jumlah *principal component*. Sebagaimana visualisasi *scatter plot* pada Gambar 6, provinsi dengan indeks tertinggi dan terendah saat $n = 3$ tetap menempati posisi ekstrem yang sama saat $n = 2$ dan $n = 4$. Selain itu, nilai *Adjusted Rand Index* (ARI) pada Gambar 5 menunjukkan hasil yang lebih moderat saat $n = 2$ dan cukup baik saat $n = 4$. Hal tersebut mengindikasikan sebagian provinsi mengalami perpindahan *cluster* ketika jumlah *principal component* diubah dari $n = 3$ menjadi $n = 2$ dan $n = 4$.

Selanjutnya, analisis sensitivitas terhadap pemilihan jumlah *cluster* dilakukan dengan membandingkan skenario jumlah *cluster* $K = 4$ dan $K = 6$ terhadap skenario awal yaitu $K = 5$. Setiap skenario pada analisis sensitivitas terhadap pemilihan jumlah *cluster* tetap menggunakan 3 *principal component* ($n = 3$). Konsistensi kelompok *cluster* dievaluasi berdasarkan *Adjusted Rand Index* (ARI) dari skenario yang dibandingkan. Detail dari hasil analisis sensitivitas tersebut terdapat pada Gambar 7 sebagai berikut:



Gambar 7. Barchart ARI pada Analisis Sensitivitas Pemilihan Jumlah Cluster

Berdasarkan Gambar 7, hasil pengujian menunjukkan bahwa penurunan jumlah *cluster* dari $K = 5$ menjadi $K = 4$ menghasilkan tingkat kesesuaian *cluster* yang relatif moderat. Temuan ini konsisten dengan hasil pada Subbab 3.4 yang menunjukkan bahwa skenario $K = 4$ menghasilkan beberapa provinsi dengan *Silhouette Coefficient* (SC) yang bernilai negatif sehingga struktur *cluster* saat $K = 4$ kurang optimal. Sebaliknya, penambahan jumlah *cluster* dari $K = 5$ menjadi $K = 6$ menghasilkan tingkat kesesuaian *cluster* yang melebihi batas ambang 0,90 yang mengindikasikan bahwa mayoritas provinsi berada dalam struktur *cluster* yang sama.

4. KESIMPULAN

Penelitian ini berhasil membangun *Social Engineering Vulnerability Index* pada 38 provinsi di Indonesia tahun 2025 dengan menggunakan *Principal Component Analysis* (PCA) yang mampu mempertahankan 79,33% variansi data serta *K-Medoids clustering* untuk menganalisis tingkatan kerentanan pada masing-masing provinsi dengan memilih $K = 5$ sebagai jumlah *cluster* optimal berdasarkan interpretasi dari *Silhouette Coefficient* (SC) dan rata-rata *Silhouette Coefficient*. Hasil penelitian menunjukkan bahwa kerentanan terhadap *social engineering* provinsi di Indonesia bervariasi cukup luas dan terbagi atas kerentanan yang sangat tinggi, tinggi, sedang, rendah, dan sangat rendah. Variasi dari hasil temuan tersebut dikarenakan adanya kesenjangan kondisi sosial-ekonomi dan ekosistem *digital* pada setiap provinsi di Indonesia dalam menghadapi *social engineering*. Selain itu, analisis sensitivitas telah dilakukan dan menunjukkan hasil yang *robust* terhadap stabilitas peringkat indeks dan bervariasi terhadap stabilitas *cluster*. Secara keseluruhan, penelitian ini mampu membangun *Social Engineering Vulnerability Index* pada tingkat subnasional yang selama ini belum banyak dikaji dalam

literatur. Penelitian ini diharapkan dapat memberikan kontribusi untuk mendukung kebijakan secara akurat dalam mengatasi kesenjangan sosial ekonomi dan ekosistem *digital* pada setiap provinsi di Indonesia dalam menghadapi serangan *social engineering*.

5. UCAPAN TERIMA KASIH

Peneliti menyampaikan ucapan terima kasih kepada Badan Pusat Statistik (BPS) Kota Bengkulu dan Program Studi S1 Matematika FMIPA Universitas Bengkulu yang telah memberikan dukungan dalam melakukan penelitian dan pembuatan artikel ini. Peneliti mengucapkan terima kasih terhadap Badan Pusat Statistik (BPS) Indonesia dan Asosiasi Penyedia Jasa Internet Indonesia (APJII) yang telah menyediakan data-data publik yang digunakan dalam penelitian. Peneliti juga mengucapkan terima kasih terhadap berbagai pihak lainnya yang telah memberikan referensi terhadap topik penelitian ini.

6. REKOMENDASI

Keterbatasan ketersediaan data historis dapat membuat nilai *Social Engineering Vulnerability Index* pada tahun berjalan belum dapat diketahui secara langsung. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan indeks ini melalui pendekatan prediksi nilai variabel sehingga hasil kerentanan dapat diperoleh lebih awal untuk mendukung pengambilan kebijakan secara tepat waktu. Kemudian, hasil indeks dalam penelitian ini belum divalidasi dengan data kejadian *social engineering* aktual di lapangan. Perlu diakui bahwa kasus *social engineering* tidak selalu dilaporkan secara resmi karena terdapat kemungkinan bahwa korban tidak menyadari serangan *social engineering* yang dialaminya. Oleh karena itu, penelitian selanjutnya disarankan untuk melakukan validasi melalui data laporan yang tercatat dan melalui survei langsung kepada masyarakat untuk menelusuri serangan *social engineering* yang tidak disadari.

7. REFERENSI

- Albladi, S. M., & Weir, G. R. S. (2020). Predicting Individuals' Vulnerability to Social Engineering in Social Networks. *Cybersecurity*, 3(1), 7. <https://doi.org/10.1186/s42400-020-00047-5>.
- Aldawood, H., & Skinner, G. (2019). Reviewing Cyber Security Social Engineering Training and Awareness Programs—Pitfalls and Ongoing Issues. *Future Internet*, 11(3), 73. <https://doi.org/10.3390/fi11030073>.
- APJII. (2025). Survei Penetrasi Internet dan Perilaku Penggunaan Internet 2025. Retrieved February 10, 2026, from APJII website: <https://survei.apjii.or.id/survei/group/11>.
- Bassett, G., Hylender, D., Langlois, P., Pinto, A., & Widup, S. (2022). 2022 Data Breach Investigations Report. Retrieved February 10, 2026, from Verizon website: <https://www.verizon.com/business/resources/reports/2022-dbir-data-breach-investigations-report.pdf>.
- BIG. (2024). Pulau Indonesia Bertambah Jadi 17.380, Mengapa Angkanya Berubah Setiap Tahun?. Retrieved February 10, 2026, from BIG website: <https://sipulau.big.go.id/news/11>.
- BPS. (2025a). Gini Ratio Menurut Provinsi dan Daerah, 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/2/OTgjMg==/gini-ratio-menurut-provinsi-dan-daerah.html>.

- BPS. (2025b). Jumlah Penduduk Pertengahan Tahun (Ribu Jiwa), 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/2/MTk3NSMy/jumlah-penduduk-pertengahan-tahun-ribu-jiwa-.html>.
- BPS. (2025c). Penduduk, Laju Pertumbuhan Penduduk, Distribusi Persentase Penduduk, Kepadatan Penduduk, Rasio Jenis Kelamin Penduduk Menurut Provinsi, 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/3/V1ZSbFRUY3lTbFpEYTNsVWNGcDZjek53YkhsNFFUMDkjMw==/penduduk--laju-pertumbuhan-penduduk--distribusi-persentase-penduduk--kepadatan-penduduk--rasio-jenis-kelamin-penduduk-menurut-provinsi--2024.html?year=2025>.
- BPS. (2025d). Persentase Penduduk Miskin (P0) Menurut Provinsi dan Daerah (Persen), 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/2/MTkyIzI=/persentase-penduduk-miskin--p0--menurut-provinsi-dan-daerah.html>.
- BPS. (2025e). Rata-Rata Lama Sekolah Penduduk Umur 15 Tahun ke Atas Menurut Provinsi, 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/2/MTQyOSMy/rata-rata-lama-sekolah-penduduk-umur-15-tahun-ke-atas-menurut-provinsi.html>.
- BPS. (2025f). Tingkat Pengangguran Terbuka Menurut Provinsi (Persen), 2025. Retrieved February 10, 2026, from BPS website: <https://www.bps.go.id/id/statistics-table/2/NTQzIzI=/tingkat-pengangguran-terbuka--februari-2024.html>.
- FIRST. (2023). Common Vulnerability Scoring System version 4.0: User Guide. Retrieved February 10, 2026, from FIRST website: <https://www.first.org/cvss/v4.0/user-guide>.
- Hossin, M. A., Du, J., Mu, L., & Asante, I. O. (2023). Big Data-Driven Public Policy Decisions: Transformation Toward Smart Governance. *Sage Open*, 13(4), 1–19. <https://doi.org/10.1177/21582440231215123>.
- Jacobs, J., Romanosky, S., Edwards, B., Adjerid, I., & Roytman, M. (2021). Exploit Prediction Scoring System (EPSS). *Digital Threats: Research and Practice*, 2(3). <https://doi.org/10.1145/3436242>.
- Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New Jersey: John Wiley & Son.
- Kumar, K., Shruthi, Sathwick, & Prathap. (2024). Outcomes of Social Engineering: Understanding Cybercriminals' Exploitation of Human Psychology. *International Journal of Scientific Research in Engineering and Management*, 08(11), 1–10. <https://doi.org/10.55041/IJSREM39133>.
- Maraj, A., & Butler, W. (2022). Taxonomy of Social Engineering Attacks: A Survey of Trends and Future Directions. *Proceedings of the 17th International Conference on Information Warfare and Security*, 17(1), 185–193. <https://doi.org/10.34190/icwsws.17.1.40>.
- Mitra, S., Nguyen, T., Zhang, Q., Kim, H., Salemi, H., Chang, C.-W., Zhang, F., Hong, M., Lu, C.-T., Purohit, H., & Cho, J.-H. (2025). SCVI: Bridging Social and Cyber Dimensions for Comprehensive Vulnerability Assessment. <http://arxiv.org/abs/2503.20806>.
- Şandor, A., Tont, G., & Simion, E. (2022). A Mathematical Model for Risk Assessment of Social Engineering Attacks. *TEM Journal*, 11(1), 334–338. <https://doi.org/10.18421/TEM111-42>.
- Trivianita, N., Marsisno, W., & Wilantika, N. (2019). Social Vulnerability Index to Tuberculosis of Provinces in Indonesia. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, 216, 170–180. <https://doi.org/10.2991/ASSDG-18.2019.15>.

- Vishwanath, A., Harrison, B., & Ng, Y. J. (2016). Suspicion, Cognition, and Automaticity Model of Phishing Susceptibility. *Communication Research*, 45(8), 1146–1166. <https://doi.org/10.1177/0093650215627483>.
- Wang, J., Herath, T., Chen, R., Vishwanath, A., & Rao, H. R. (2012). Research Article Phishing Susceptibility: An Investigation Into the Processing of a Targeted Spear Phishing Email. *IEEE Transactions on Professional Communication*, 55(4), 345–362. <https://doi.org/10.1109/TPC.2012.2208392>.