

Analisis Faktor-Faktor Tingkat Penggunaan *E-Commerce* di Setiap Provinsi Indonesia Menggunakan Analisis Faktor

Farhan Muhammad Rizqi¹, Muhammad Rafli Nugrahasyach², Sri Pingit Wulandari³

¹²³Program Studi Sarjana Terapan, Departemen Statistika Bisnis, Fakultas Vokasi, Institut Teknologi Sepuluh Nopember, Surabaya

ARTICLE INFO

Article history:

Received Oktober 27, 2024

Revised Oktober 29, 2024

Accepted November 15, 2024

Available online 19 November, 2024

Keywords:

Analisis faktor, principal component analysis, tingkat penggunaan e-commerce

Keywords:

Factor analysis, principal component analysis, level of e-commerce usage



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2024 by Author. Published by Yayasan Daarul Huda

ABSTRAK

Peningkatan persentase penggunaan *e-commerce* di setiap provinsi di Indonesia dipengaruhi oleh sejumlah faktor sosioekonomi dan digital yang beragam. Studi ini bertujuan untuk menganalisis faktor-faktor utama yang memengaruhi tingkat penggunaan *e-commerce* di Indonesia pada tahun 2023. Penelitian dilakukan dengan menggunakan analisis faktor dari 10 variabel, yaitu Rasio Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), Tingkat Partisipasi Tenaga Kerja (TPT), Indeks Literasi Digital, Teknologi Informasi dan Komunikasi (TIK), proporsi pengguna internet, serta beberapa indikator Indeks Masyarakat Digital (IMD) seperti infrastruktur dan ekosistem digital, keterampilan digital, pemberdayaan, dan kesempatan kerja berbasis teknologi. Analisis faktor dengan metode *Principal Component Analysis* (PCA) digunakan untuk mengidentifikasi variabel-variabel dominan yang terkait dengan adopsi *e-commerce*, terutama di 34 provinsi di Indonesia. Hasil analisis menunjukkan bahwa variabel yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi relatif homogen. Uji asumsi pada data menunjukkan terpenuhinya asumsi distribusi normal multivariat, dependen, dan kecocokan variabel untuk difaktorkan. Dari PCA, diidentifikasi dua faktor utama yang menjelaskan 61,237% variabilitas data: Faktor pertama adalah Pendidikan dan Kualitas Sumber Daya Manusia, yang meliputi variabel Rata-rata Lama Sekolah, Tingkat Pengangguran Terbuka, Indeks Literasi Digital, Teknologi Informasi dan Komunikasi, serta Pemberdayaan. Faktor kedua adalah Keterampilan dan Ketenagakerjaan, yang mencakup Keterampilan Digital dan Pekerjaan berbasis teknologi.

ABSTRACT

The increase in the percentage of *e-commerce* use in each province in Indonesia is influenced by a number of diverse socioeconomic and digital factors. This study aims to analyze the main factors that influence the level of *e-commerce* use in Indonesia in 2023. The study was conducted using factor analysis of 10 variables, namely the Schooling Years Ratio (RLS), Expected Years of Schooling (HLS), Labor Force Participation Rate (TPT), Digital Literacy Index, Information and Communication Technology (ICT), the proportion of internet users, and several indicators of the Digital Society Index (IMD) such as digital infrastructure and ecosystem, digital skills, empowerment, and technology-based employment opportunities. Factor analysis using the Principal Component Analysis (PCA) method was used to identify the dominant variables related to *e-commerce* adoption, especially in 34 provinces in Indonesia. The results of the analysis show that the variables that influence the percentage of *e-commerce* use in each province are relatively homogeneous. The assumption test on the data shows that the assumptions of multivariate normal distribution, dependent, and the suitability of the variables for factoring are met. From PCA, two main factors were identified that explained 61.237% of the data variability: The first factor is Education and Human Resource Quality, which includes the variables Average Length of Schooling, Open Unemployment Rate, Digital Literacy Index, Information and Communication Technology, and Empowerment. The second factor is Skills and Employment, which includes Digital Skills and Technology-based Jobs.

PENDAHULUAN

Dalam era digital yang semakin maju, *e-commerce* telah menjadi komponen penting dalam mendorong pertumbuhan ekonomi nasional. Dengan nilai transaksi yang mencapai Rp 533 triliun pada tahun 2023, sektor ini menunjukkan potensi besar dalam membangun ekonomi digital, khususnya bagi pelaku UMKM (Kementerian Perdagangan RI, 2024). *E-commerce* tidak hanya berfungsi sebagai sarana transaksi daring, tetapi juga berperan sebagai penggerak ekonomi yang memberdayakan UMKM untuk bersaing di pasar digital. Namun, literasi digital masyarakat Indonesia masih rendah dan keterbatasan

*Hernan Pratama

E-mail addresses: raflinugrahasyach26@gmail.com

infrastruktur Teknologi Informasi dan Komunikasi (TIK) di beberapa wilayah menjadi hambatan dalam meningkatkan penggunaan *e-commerce*. Walaupun akses internet terus meningkat, pemanfaatan teknologi untuk mendukung ekonomi digital masih terbatas, terutama di kalangan masyarakat dengan tingkat literasi digital yang rendah. Faktor-faktor seperti tingkat pendidikan, rata-rata lama sekolah, dan indeks literasi digital sangat berpengaruh terhadap kesiapan masyarakat dalam memanfaatkan *e-commerce* sebagai bagian dari aktivitas ekonomi harian mereka (Katadata, 2024).

Untuk mengidentifikasi faktor-faktor utama dalam penggunaan *e-commerce* di Indonesia, pendekatan *Principal Component Analysis* (PCA) diperlukan. Analisis ini menyederhanakan berbagai variabel terkait sehingga faktor-faktor dominan yang memengaruhi penggunaan *e-commerce* di setiap provinsi dapat diidentifikasi. *Factor analysis* adalah metode yang mengelompokkan variabel dengan korelasi tinggi ke dalam faktor-faktor utama yang lebih sederhana, sangat relevan untuk mengetahui tingkat penggunaan *e-commerce* di Indonesia. Berbagai indikator seperti tingkat literasi digital, penetrasi internet, dan infrastruktur TIK yang kompleks dapat dianalisis secara lebih efisien dengan pendekatan ini (Hair, 2019).

Dengan PCA, analisis faktor dapat mereduksi dimensi data berkorelasi menjadi komponen utama yang independen, membantu mengidentifikasi faktor dominan seperti akses teknologi, keterampilan digital, dan penyediaan infrastruktur. Hasil dari penelitian ini diharapkan mampu menjadi landasan kebijakan yang efektif untuk memperluas penggunaan *e-commerce*. Temuan penelitian ini juga diharapkan dapat memberikan rekomendasi ilmiah yang memperkuat kebijakan strategis untuk meningkatkan penggunaan *e-commerce* di seluruh wilayah Indonesia.

TINJAUAN PUSTAKA

Statistika Deskriptif

Statistika deskriptif adalah statistik yang digunakan untuk menganalisis data yang telah terkumpul sebagaimana adanya tanpa bermaksud membuat kesimpulan yang telah berlaku untuk umum atau generasi (Yusuf, 2016). Data yang telah dikumpulkan nantinya akan dianalisis dan kemudian akan disajikan dengan baik dalam bentuk *bar chart*, *pie chart*, dan histogram, maupun yang lainnya sehingga para pembaca dapat lebih mudah mengerti.

Mean

Mean adalah nilai rata-rata sebuah data. Nilai rata-rata merupakan total penjumlahan dibagi dengan jumlah data. Nilai rata-rata mewakili nilai dari sekumpulan data yang dianggap sebagai nilai yang paling dekat dengan ukuran sebenarnya (Ismail, 2018). Rumus mencari *mean* dapat dilihat pada persamaan 2.1 dan 2.2 sebagai berikut.

Rumus *Mean* Data Tunggal:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (2.1)$$

Rumus *Mean* Data Kelompok:

$$\bar{X} = \frac{\sum_{i=1}^n X_i f_i}{\sum f_i} \quad (2.2)$$

Keterangan :

- $\bar{X}$ = Nilai rata-rata
- X_i = Nilai data ke-i
- n = Banyak data
- f_i = Frekuensi kelompok data ke-i

Standar Deviasi

Standar deviasi adalah ukuran statistik yang menunjukkan seberapa tersebar atau bervariasi data dalam suatu kumpulan. Standar deviasi menggambarkan seberapa jauh rata-rata nilai dari data tersebut menyimpang dari nilai rata-ratanya. Secara matematis, standar deviasi dihitung sebagai akar kuadrat dari varians, di mana varians adalah rata-rata dari kuadrat selisih antara setiap nilai data dengan rata-rata keseluruhan data (Setyo, 2019). Rumus standar deviasi ditunjukkan pada Persamaan 2.3 sebagai berikut.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (2.3)$$

Keterangan :

- $\bar{X}$ = Nilai rata-rata

X_i = Nilai data ke-i
 n = Banyak data
 σ = Standar deviasi

Asumsi Analisis Faktor

Analisis faktor didasarkan pada beberapa asumsi, yaitu asumsi distribusi normal multivariat, kecukupan data yang ditentukan oleh nilai Kaiser-Meyer-Olkin lebih dari 0,5, uji keterkaitan antar variabel menggunakan uji Bartlett, serta korelasi antar variabel yang memenuhi nilai Measure of Sampling Adequacy (MSA) lebih dari 0,5 (Tabachnick, 2019).

Uji Distribusi Normal Multivariat

Uji asumsi distribusi normal multivariat bertujuan untuk menentukan apakah suatu dataset multivariat mengikuti distribusi normal. Uji distribusi normal multivariat dilakukan menggunakan uji t proporsi, di mana dataset dianggap berdistribusi normal multivariat jika nilai proporsinya berada dalam rentang 45–55%. Nilai t proporsi diperoleh melalui korelasi antara nilai d_j^2 dengan $q_{c,p}$ (Mishra, 2022). Adapun hipotesis yang diajukan dalam pengujian ini adalah sebagai berikut.

H_0 : Data berdistribusi normal multivariat

H_1 : Data tidak berdistribusi normal multivariat

Taraf signifikansi : $\alpha = 0,05$

Daerah Penolakan : Tolak H_0 jika $45\% \leq T \text{ proporsi} \leq 55\%$ dan $p\text{-value} < \alpha = 0,05$

Statistik uji : $d_j^2 = (x_j - \bar{x})^T S^{-1} (x_j - \bar{x})$ (2.5)

Keterangan :

x_j = Objek pengamatan ke - j

$\bar{x}$ = Vektor rata - rata variabel

S = matriks kovarians

Pemeriksaan Nilai KMO

Pemeriksaan kecukupan data *Kaiser-Meyer-Olkin* (KMO) bertujuan untuk memastikan bahwa data yang tersedia cukup memadai untuk dianalisis menggunakan metode faktor (Cerny, 2022). Rumus untuk memeriksa kecukupan data KMO ditunjukkan pada Persamaan 2.6.

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j a_{ij}^2} \quad (2.6)$$

Keterangan :

r_{ij} = Koefisien korelasi

a_{ij} = Koefisien korelasi parsial

Nilai KMO sendiri memiliki rentang 0 sampai dengan 1 dengan masing-masing kriteria rekomendasi ditunjukkan pada Tabel 2.1.

Tabel 1 Kriteria Nilai KMO

Nilai KMO	Kriteria Rekomendasi
$\geq 0,9$	Sangat Baik
0,8 – 0,89	Berguna
0,70 – 0,79	Biasa
0,60 – 0,69	Cukup
0,50 – 0,59	Buruk
$< 0,50$	Tidak diterima

Uji Independensi (Uji Barlett)

Variabel dianggap saling independen jika matriks korelasi antar variabel membentuk matriks identitas. Uji *Bartlett* digunakan untuk menguji kebebasan antar variabel ini (Koot, 2019). Adapun hipotesis dan statistik uji *Bartlett* akan dijelaskan sebagai berikut.

Hipotesis:

$H_0 : \rho = I$ (Matriks korelasi identik dengan matrik identitas atau antar variabel independen)

$H_1 : \rho \neq I$ (Matriks korelasi tidak identik dengan matrik identitas atau antar variabel dependen)

Daerah penolakan : Tolak H_0 jika $\chi_{hit}^2 > \chi_{\alpha;df}^2$ atau $P\text{-Value} < \alpha$

$$\text{Statistik Uji} : \chi^2_{hit} = - \left(n - 1 - \frac{2p + 5}{6} \right) \ln |R| \quad (2.7)$$

Keterangan :

n = Banyak sampel

p = Jumlah variabel dependen

R = Nilai korelasi

Uji Korelasi Anti-Image

Matriks *Anti-Image* digunakan untuk mengidentifikasi variabel-variabel yang layak dianalisis faktor serta untuk menentukan apakah faktor-faktor yang digunakan memiliki korelasi yang cukup kuat, yaitu dengan nilai sama atau lebih dari 0,5. Jika nilai tersebut tercapai, semua faktor pembentuk variabel dianggap valid tanpa adanya reduksi faktor. Dalam matriks *Anti-Image*, variabel pertama yang harus dieliminasi adalah yang memiliki *Measure of Sampling Adequacy* (MSA) di bawah 0,5. Hasil perhitungan korelasi anti-image ditunjukkan dalam bentuk tabel, dengan angka-angka di diagonal bertanda "a" yang menunjukkan nilai MSA suatu variabel. Jika MSA variabel berada di bawah 0,5, variabel tersebut harus dikeluarkan dan proses seleksi variabel perlu diulang (Beavers, 2014). Nilai MSA berkisar dari 0 hingga 1, dengan kriteria sebagai berikut.

Tabel .1 Kriteria Nilai MSA

Nilai MSA	Kriteria
a = 1	Variabel dapat diprediksi tanpa kesalahan oleh variabel lain
$0,5 \leq a < 1$	Variabel dapat diprediksi dan dianalisis lebih lanjut
$a < 0,5$	Variabel tidak dapat diprediksi dan tidak dapat dianalisis lebih lanjut

Analisis Faktor dengan Principal Component Analyst (PCA)

Analisis faktor adalah serangkaian prosedur matematis kompleks yang digunakan untuk mengidentifikasi dan menjelaskan keterkaitan antar variabel dengan mengelompokkan variabel-variabel tersebut ke dalam sejumlah kelompok terbatas yang disebut faktor. Analisis faktor dapat dilakukan menggunakan *Principal Component Analysis* (PCA). PCA, atau analisis komponen utama, adalah teknik yang menyederhanakan data melalui transformasi linier, menghasilkan sistem koordinat baru dengan varians maksimum (Kaiser, 2020). Tujuan PCA adalah mereduksi dimensi data menjadi lebih rendah dengan sedikit kehilangan informasi. Beberapa hal yang perlu diperhatikan dalam analisis faktor adalah sebagai berikut.

Nilai Communalities

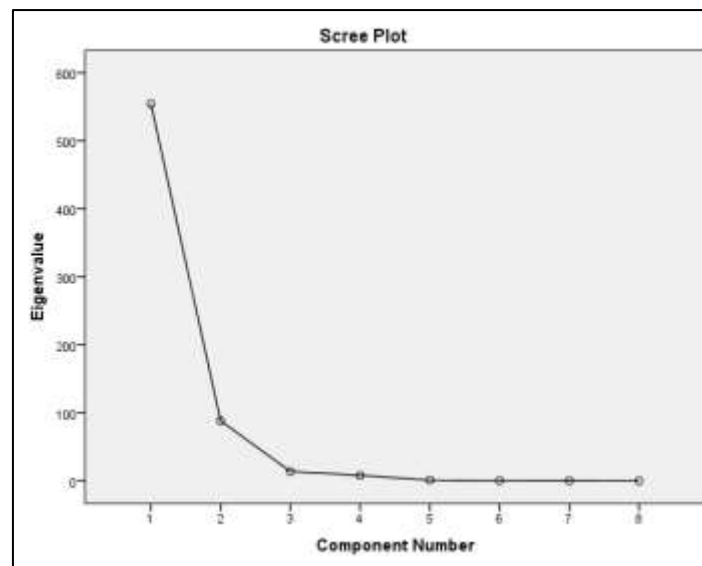
Nilai communalities mencerminkan sejauh mana keragaman atau variasi dari variabel asal dapat dijelaskan oleh faktor-faktor yang terbentuk melalui analisis faktor. Nilai yang rendah ($< 0,3$) menunjukkan bahwa variabel tersebut memiliki kontribusi yang rendah dalam model dan mungkin kurang sesuai atau sejalan dengan variabel lainnya dalam kelompok faktor yang ada. Karena itu, variabel dengan nilai communalities di bawah 0,3 umumnya disarankan untuk dikeluarkan dari analisis untuk menjaga kualitas faktor yang dihasilkan (Bandalos, 2020).

Total Variance Explained

Total varians yang dijelaskan (*total variance explained*) merujuk pada varians yang dapat diterangkan oleh komponen yang diekstraksi dan komponen yang diputar dalam analisis faktor. Bagian awal dari tabel total variance explained menunjukkan nilai eigen untuk setiap komponen. Nilai eigen ini merupakan total varians yang dijelaskan oleh setiap komponen pada variabel asli dan dihitung sebagai jumlah kuadrat dari faktor loading untuk setiap variabel yang berhubungan dengan komponen tersebut. Faktor *loading* adalah koefisien yang menunjukkan korelasi antara variabel dan faktor, di mana komponen dengan nilai eigen lebih dari satu umumnya dianggap signifikan dalam kontribusinya terhadap penjelasan varians total. Persentase varians menunjukkan seberapa besar bagian varians total yang dijelaskan oleh setiap komponen relatif terhadap keseluruhan varians pada semua variabel dalam model (Field, 2021).

Scree Plot

Scree plot adalah salah satu metode visualisasi dalam analisis komponen utama yang berfungsi untuk menampilkan grafik dari *total variance explained*. Grafik ini menunjukkan perubahan atau transisi dari satu faktor ke faktor lainnya melalui garis menurun pada sumbu y, yang mengindikasikan nilai eigen (*eigenvalues*) untuk masing-masing faktor yang terbentuk. Pada *scree plot*, sumbu x merepresentasikan jumlah komponen atau faktor yang dihasilkan, sedangkan sumbu y menunjukkan nilai eigen untuk masing-masing faktor tersebut. Grafik ini membantu dalam menentukan jumlah faktor optimal yang dapat digunakan dengan mencari "*elbow point*," yaitu titik di mana garis grafik mulai mendatar, menunjukkan faktor tambahan yang memiliki kontribusi minimal terhadap total varians (Hwang, 2022).



Gambar 1 Scree Plot

Total Variance Rotation

Tujuan utama dalam proses rotasi pada analisis faktor adalah mencapai kesederhanaan struktur faktor sehingga interpretasi dapat dilakukan dengan lebih mudah dan jelas. Dalam rotasi ortogonal, sumbu faktor dipertahankan tegak lurus satu sama lain, menghasilkan faktor yang tetap independen satu sama lain. Teknik ini sering digunakan ketika tujuan analisis adalah mereduksi jumlah variabel sambil menjaga masing-masing faktor tetap terpisah tanpa mempertimbangkan hubungan antar faktor. Dalam analisis faktor, metode rotasi ortogonal seperti Varimax memungkinkan peningkatan kejelasan struktur faktor dan membantu mengidentifikasi variabel-variabel yang memiliki pengaruh besar pada faktor tertentu, tanpa adanya korelasi antar faktor yang terbentuk (Floyd, 2022).

Pengelompokkan Faktor Baru

Matriks *rotated component* menunjukkan bagaimana variabel-variabel berkorelasi dengan komponen faktor yang terbentuk. Tabel ini mengategorikan variabel-variabel berdasarkan korelasinya dengan faktor-faktor spesifik, sehingga nilai *factor loading* yang tinggi menunjukkan bahwa suatu variabel memiliki korelasi kuat dengan faktor tersebut. Nilai *loading* yang lebih tinggi, umumnya di atas 0,5, menunjukkan bahwa variabel tersebut sangat berkaitan dengan faktor yang bersangkutan, sehingga mempermudah interpretasi dan identifikasi kelompok variabel yang berkontribusi besar pada komponen faktor tertentu (Cohen, 2020).

Pemberian Nama Faktor Baru

Penamaan faktor dilakukan dengan mengidentifikasi esensi yang mendasari variabel-variabel dalam satu faktor, dan memberi nama yang dapat mencerminkan karakteristik dominan dari variabel-variabel tersebut. Langkah ini melibatkan proses generalisasi untuk menggambarkan keterkaitan antara variabel-variabel awal yang berkumpul dalam satu faktor, sehingga nama faktor menjadi representatif terhadap kumpulan variabel tersebut. Pendekatan ini memungkinkan faktor yang terbentuk untuk lebih mudah dipahami dan relevan dengan tujuan analisis (Rummel, 2021).

RLS (Rata-rata Lama Sekolah)

RLS adalah ukuran rata-rata tahun pendidikan yang diselesaikan oleh individu berusia 15 tahun ke atas di suatu wilayah. RLS mencerminkan akses dan pencapaian pendidikan masyarakat, serta berperan sebagai indikator kualitas pendidikan di suatu wilayah. Meningkatnya RLS umumnya berhubungan dengan kualitas sumber daya manusia yang lebih baik (Badan Pusat Statistik, 2023).

HLS (Harapan Lama Sekolah)

HLS menggambarkan jumlah tahun pendidikan yang diharapkan akan diselesaikan oleh anak-anak yang baru memasuki sistem pendidikan. HLS memberikan indikasi terhadap tingkat pendidikan di masa depan, serta ekspektasi masyarakat dalam memperoleh pendidikan tinggi. Tingginya HLS mencerminkan harapan besar terhadap akses pendidikan jangka panjang (Badan Pusat Statistik, 2023).

TPT (Tingkat Pengangguran Terbuka)

TPT adalah persentase dari angkatan kerja yang aktif mencari pekerjaan namun belum memperoleh pekerjaan. Tingginya TPT mengindikasikan permasalahan ekonomi yang dapat mempengaruhi kesejahteraan masyarakat dan stabilitas ekonomi suatu daerah (Badan Pusat Statistik, 2023).

Indeks Literasi Digital

Indeks Literasi Digital mengukur kemampuan individu dalam menggunakan teknologi digital secara bijak, aman, dan produktif. Indeks ini mencakup keterampilan dalam menggunakan perangkat digital, memahami etika di ruang digital, keamanan siber, dan budaya digital (Badan Pusat Statistik, 2023).

TIK (Teknologi Infomasi dan Komunikasi)

TIK mencakup semua teknologi yang digunakan untuk mengelola dan menyampaikan informasi melalui perangkat digital, termasuk internet, komputer, dan perangkat mobile. Tingkat penggunaan TIK mencerminkan kemajuan suatu negara dalam adopsi teknologi modern untuk mendukung berbagai sektor, termasuk pendidikan, ekonomi, dan pemerintahan (Pew Research Center, 2022).

Proporsi Pengguna Internet

Variabel ini menunjukkan persentase populasi yang menggunakan internet dalam kehidupan sehari-hari. Tingkat penggunaan internet dapat menggambarkan akses terhadap informasi, literasi digital, dan kesenjangan digital antarwilayah, terutama di negara berkembang (World Economic Forum, 2022).

Infrastruktur dan Ekosistem (IMD)

Indikator Infrastruktur dan Ekosistem dalam IMD menggambarkan kesiapan infrastruktur digital dan ekosistem pendukung, termasuk jaringan internet, perangkat digital, dan kebijakan yang mendukung transformasi digital. Infrastruktur yang kuat mendukung akses luas terhadap teknologi dan literasi digital (OECD, 2021).

Pemberdayaan (IMD)

Indikator pemberdayaan dalam IMD mengukur tingkat kesiapan masyarakat dalam menggunakan teknologi digital untuk kegiatan produktif, seperti *e-commerce*, pembelajaran digital, dan keterlibatan sosial. Tingkat pemberdayaan digital yang tinggi menunjukkan kemampuan masyarakat memanfaatkan teknologi untuk kegiatan ekonomi dan pendidikan (Bertot, 2020).

Pekerjaan (IMD)

Pekerjaan dalam IMD mencakup akses terhadap pekerjaan yang didukung teknologi digital, yang juga meliputi peluang ekonomi di industri digital dan kemajuan karier berbasis teknologi. Faktor ini berfokus pada kontribusi teknologi dalam menciptakan lapangan pekerjaan baru dan meningkatkan keterampilan tenaga kerja (McKinsey Global Institute, 2021).

METODE

Sumber Data

Sumber data yang digunakan pada praktikum ini adalah data sekunder yang diperoleh dari berbagai sumber, yaitu Badan Pusat Statistik (BPS) yang meliputi rata-rata lama sekolah, harapan lama sekolah, tingkat pengangguran terbuka, dan indeks teknologi informasi dan komunikasi, APJII (Asosiasi Penyelenggara Jasa Internet Indonesia) yang meliputi proporsi pengguna internet, dan KemenKomDigi (Kementerian Komunikasi dan Digital) yang mencakup indeks literasi digital dan empat subpilar dari indeks masyarakat digital yaitu infrastruktur ekosistem, keterampilan digital, pemberdayaan, dan pekerjaan. Data yang digunakan sebanyak 34 data dengan unit observasi yaitu provinsi di Indonesia. Pengambilan data dilakukan pada Hari Rabu, 23 Oktober 2024 pukul 08.00 WIB di Kota Surabaya, Jawa Timur.

Variabel Penelitian

Variabel penelitian yang digunakan dalam praktikum ini ditunjukkan pada Tabel berikut.

Tabel 2 Variabel Penelitian

Variabel	Keterangan	Skala Pengukuran	Satuan
X₁	Rata-rata Lama Sekolah (RLS)	Rasio	Tahun
X₂	Harapan Lama Sekolah (HLS)	Rasio	Tahun
X₃	Tingkat Pengangguran Terbuka (TPT)	Rasio	Persen
X₄	Indeks Literasi Digital	Interval	-
X₅	Teknologi Informasi dan Komunikasi (TIK)	Interval	-
X₆	Proporsi Pengguna Internet	Interval	Persen
X₇	Infrastruktur Ekosistem (IMD)	Interval	-
X₈	Keterampilan Digital (IMD)	Interval	-
X₉	Pemberdayaan (IMD)	Interval	-
X₁₀	Pekerjaan (IMD)	Interval	-

Struktur Data

Struktur data yang digunakan dalam praktikum ini ditunjukkan Tabel 3.2.

Tabel 3. Struktur Data

Data ke-	X ₁	X ₂	X ₃	...	X ₁₀
1	X ₁₁	X ₂₁	X ₃₁	...	X _{10 1}
2	X ₁₂	X ₂₂	X ₃₂	...	X _{10 2}
3	X ₁₃	X ₂₃	X ₃₃	...	X _{10 3}
⋮	⋮	⋮	⋮	...	⋮
34	X _{1 34}	X _{2 34}	X _{3 34}	...	X _{10 34}

Langkah Analisis

Langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut.

1. Mengumpulkan data sekunder faktor-faktor yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023.
2. Mendeskripsikan karakteristik data.
3. Melakukan pemeriksaan dan pengujian asumsi analisis faktor yang terdiri atas empat pengujian maupun pemeriksaan sebagai berikut.
 - a. Uji distribusi normal multivariat menggunakan uji korelasi antara jarak mahalanobis dengan nilai kuantil.
 - b. Uji korelasi dependensi menggunakan uji *Bartlett*.
 - c. Pemeriksaan kecukupan data menggunakan uji KMO.
 - d. Pemeriksaan korelasi antar variabel menggunakan *anti image correlation*.
4. Melakukan analisis faktor dengan tahapan sebagai berikut.
 - a. Menentukan jumlah komponen.
 - b. Mengelompokkan dan memberi nama faktor yang terbentuk.
 - c. Pemeriksaan ketepatan komponen
5. Menginterpretasikan hasil analisis.
6. Menarik kesimpulan dan saran.

HASIL DAN PEMBAHASAN

Bab ini akan membahas terkait dengan analisis faktor pada faktor-faktor yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi di Indonesia tahun 2023 yang meliputi karakteristik data, pemeriksaan asumsi analisis faktor, dan analisis faktor dengan *principal component analysis*.

Karakteristik Data

Karakteristik data faktor-faktor yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 disajikan pada Tabel 4.1.

Tabel 3 Karakteristik Data

Variabel	Mean	Standar Deviasi	Minimum		Maksimum	
			Nilai	Provinsi	Nilai	Provinsi
X ₁	8,93	0,91	7,15	Papua	11,45	DKJ
X ₂	13,30	0,74	11,15	Papua	15,66	DIY
X ₃	4,61	1,42	2,27	Sulawesi Barat	7,52	Banten
X ₄	3,64	0,10	3,41	Maluku Utara	3,88	DKJ
X ₅	76,44	10,74	33,84	Papua	93,98	DKJ
X ₆	67,43	8,19	37,95	Papua	82,47	DKJ
X ₇	57,87	4,49	50,35	Sulawesi Barat	72,68	DKJ
X ₈	57,37	3,55	48,92	Kalimantan Utara	64,72	Bengkulu
X ₉	26,49	1,85	22,75	NTT	31,42	DKJ
X ₁₀	31,94	4,81	16,66	NTT	39,79	Bali

Tabel 3 menunjukkan bahwa nilai standar deviasi pada masing-masing variabel yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 bernilai lebih kecil daripada nilai rata-ratanya, sehingga dapat dikatakan data kurang bervariasi atau homogen. Tabel 3 juga menunjukkan nilai minimum dan maksimum dari setiap variabel prediktor dari 34 provinsi di Indonesia. Provinsi-provinsi dengan nilai terendah pada variabel-variabel tersebut, seperti Papua yang memiliki Rata-rata Lama Sekolah (RLS) terendah sebesar 7,15 tahun dan Proporsi Pengguna Internet terendah sebesar 37,95 persen, serta Nusa Tenggara Timur (NTT) yang memiliki nilai pemberdayaan dan

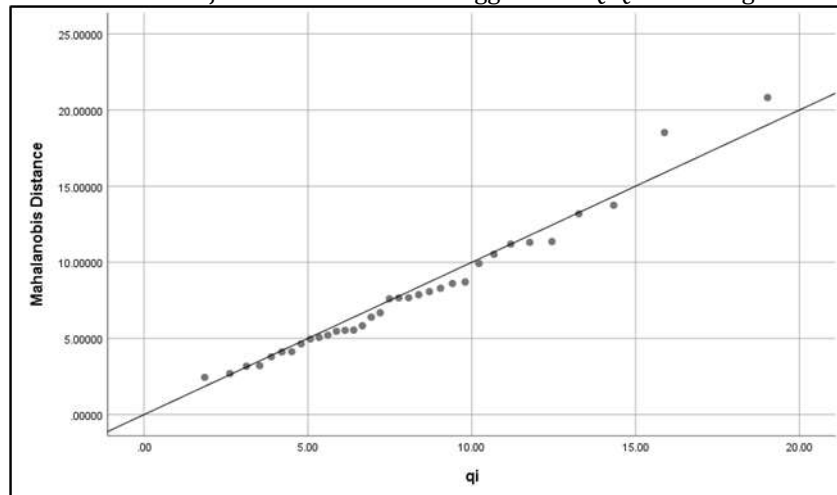
pekerjaan dalam ekosistem digital terendah yang mencerminkan adanya ketimpangan akses terhadap pendidikan, infrastruktur digital, dan pemberdayaan ekonomi digital di berbagai wilayah Indonesia.

Pemeriksaan Asumsi Analisis Faktor pada Data Faktor-faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

Pengujian asumsi dilakukan sebelum melakukan analisis faktor dengan menggunakan analisis komponen utama pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 sebagai berikut.

Pengujian Asumsi Distribusi Normal Multivariat Faktor-Faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

Uji distribusi normal multivariat dilakukan untuk menguji apakah data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023. Hasil pemeriksaan distribusi normal multivariat disajikan secara visual menggunakan *Q-Q Plot* sebagai berikut.



Gambar 3 Q-Q Plot Distribusi Normal Multivariat

Berdasarkan Gambar 4.1 menunjukkan bahwa *Q-Q Plot* antara nilai kuantil dengan jarak *Mahalanobis* plot-plotnya mengikuti garis linear. Maka, dapat disimpulkan bahwa secara visual data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 berdistribusi normal multivariat. Kemudian, dilakukan pengujian distribusi normal secara statistik menggunakan *T-Proporsi* untuk memperkuat hasil pemeriksaan secara visual pada faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023. Hipotesis uji *T-Proporsi* data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 adalah sebagai berikut.

Hipotesis :

H_0 : Data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 berdistribusi normal multivariat

H_1 : Data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 tidak berdistribusi normal multivariat

Diketahui taraf signifikan (α) sebesar 0,05 dengan daerah penolakan tolak jika *T-Proporsi* di luar rentang $45\% \leq \alpha \leq 55\%$. Didapatkan bahwa nilai *T-proporsi* sebesar 0,500 atau 50% di mana nilai tersebut berada di dalam $45\% \leq \alpha \leq 55\%$, maka diputuskan gagal tolak H_0 yang artinya data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 berdistribusi normal multivariat

Uji Bartlett Faktor-faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

Uji *bartlett* digunakan untuk memeriksa apakah data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 sebagai berikut.

Hipotesis :

H_0 : $\rho = I$ (Data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 antar variabel independen)

H_1 : $\rho \neq I$ (Data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 antar variabel dependen)

Taraf signifikan yang digunakan pada penelitian ini adalah α sebesar 0,05 dan tolak H_0 jika $\chi^2 > \chi^2_{(0,05;28)}$ atau *p-value* < α . Statistik uji *bartlett* disajikan pada tabel 4. berikut.

Tabel 4. Uji Distribusi *Bartlett*

χ^2	$\chi^2_{(0,05;28)}$	P-Value
132,689	41,337	<0,01

Berdasarkan Tabel 4 diperoleh bahwa sebesar 132,689 yang lebih kecil dari nilai α sebesar 41,337 atau p-value yang lebih kecil dari 0,05 yang lebih besar dari α sebesar 0,05 maka diputuskan tolak yang artinya Data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 antar variabel dependen.

Uji Kaiser Meyer Olkin (KMO) Faktor-faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

Pemeriksaan KMO digunakan untuk menunjukkan suatu metode sampling untuk mengukur kecukupan sampel yang digunakan sudah memenuhi syarat atau tidak. Uji KMO pada faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 yang diperoleh hasil nilai KMO sebesar 0,685 yang berada di dalam selang $0,6 < KMO < 0,69$ yang artinya data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 sudah cukup untuk difaktorkan dan termasuk dalam kriteria data cukup.

Uji Korelasi Anti-Image Faktor-Faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

Korelasi *Anti-Image* melalui nilai MSA dilakukan untuk mengukur kecukupan sampling dari setiap variabel. Syarat diterimanya korelasi *Anti-Image* yaitu apabila nilai MSA berada di atas 0,5, maka variabel tersebut dapat diprediksi dan dianalisis lebih lanjut. Sedangkan, apabila nilai MSA yang didapatkan berada di bawah 0,5, maka variabel tersebut tidak dapat diprediksi dan dianalisis lebih lanjut sehingga variabel tersebut harus dieliminasi. Uji korelasi *Anti-Image* pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 sebagai berikut.

Tabel 5 Uji Korelasi *Anti-Image*

Variabel	MSA
RLS	0,761
TPT	0,730
Indeks Literasi Digital	0,668
TIK	0,632
Internet	0,619
Keterampilan Digital	0,639
Pemberdayaan	0,801
Pekerjaan	0,659

Berdasarkan Tabel 4.3 menunjukkan bahwa hasil uji korelasi *Anti-Image* diperoleh nilai korelasi pada setiap variabel lebih besar dari 0,5 yang artinya variabel yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 dapat diprediksi dan dianalisis lebih lanjut.

Analisis Faktor dengan Principal Component Analysis pada Data Faktor-faktor yang Mempengaruhi Persentase Penggunaan E-Commerce di Setiap Provinsi Indonesia Tahun 2023

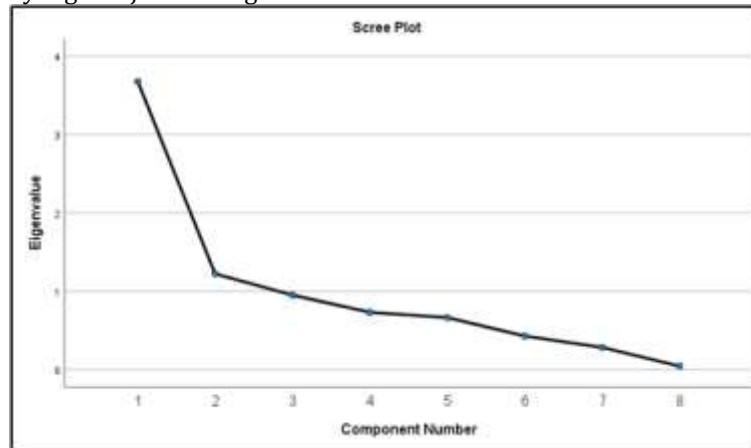
Principal Component Analysis dilakukan untuk mengetahui jumlah variabel yang dapat mewakili kondisi keragaman variabel yang diuji, sehingga dapat diduga jumlah faktor baru yang lebih sedikit yang akan terbentuk. Salah satu cara menentukan banyaknya *principal component* yaitu dengan mencari *eigen value* yang lebih besar dari satu dan secara visual dapat dilihat menggunakan *scree-plot*. Nilai *eigen value* pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 sebagai berikut.

Tabel 6 Nilai *Eigen Value*

Komponen	<i>Eigen Value</i>		
	Total	Varians (%)	Kumulatif (%)
1	3,676	45,954	45,954
2	1,223	15,283	61,237
3	0,950	11,881	73,118
4	0,731	9,132	82,250
5	0,664	8,300	90,551

6	0,428	5,354	95,905
7	0,283	3,542	99,447
8	0,044	0,553	100.000

Berdasarkan Tabel 4.4 Menunjukkan bahwa terdapat 2 komponen yang memiliki nilai eigen value lebih dari satu. Komponen 1 memiliki nilai *eigen value* sebesar 3,676 dengan persen keragaman sebesar 45,954% yang artinya komponen 1 mampu menjelaskan variabel sebesar 45,954%. Komponen 2 memiliki nilai *eigen value* sebesar 1,223 dengan persen keragaman sebesar 15,283% yang artinya komponen 2 mampu menjelaskan variabel sebesar 15,283%. Secara keseluruhan dengan 2 faktor saja dapat menjelaskan 61,237% keragaman yang dimiliki oleh variabel pada faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023. Secara visual dapat dilihat dari *screeplot* yang disajikan sebagai berikut.



Gambar 4 *Scree Plot* Distribusi Normal Multivariat

Berdasarkan Gambar 4 menunjukkan secara visual terjadi kondisi turun drastis dari titik komponen 1 ke titik komponen 2. Kemudian kondisi penurunan dari titik komponen 2 ke titik 4 terlihat cukup signifikan. Namun, setelah melewati titik komponen ke 4 garis kurva naik sedikit pada titik komponen 5 lalu mulai melandai, sampai pada titik ke 8 penurunan yang terjadi semakin landai. Hal ini didukung oleh titik 1 hingga titik 2 mempunyai eigen value lebih dari 1, sedangkan faktor lainnya memiliki eigen value kurang dari 1, maka dengan 2 faktor saja sudah dikatakan efisien untuk menjelaskan variabel secara keseluruhan.

Analisis komponen utama pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 terdiri dari nilai *communalities*, total varians, scree plot, dan pengelompokan komponen. Nilai *comumunalities* pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 adalah sebagai berikut.

Tabel 7 Nilai *Communalities*

Variabel	Initial	Extraction
RLS	1,000	0,729
TPT	1,000	0,450
Indeks Literasi Digital	1,000	0,218
TIK	1,000	0,837
Internet	1,000	0,821
Keterampilan Digital	1,000	0,596
Pemberdayaan	1,000	0,633
Pekerjaan	1,000	0,616

Berdasarkan Tabel 7 menunjukkan bahwa dari 8 variabel yang digunakan, terdapat 6 variabel memiliki nilai *comumunalities* sebesar lebih dari 0,5 yang artinya data tersebut yaitu RLS, TIK, Internet, Keterampilan digital, Pemberdayaan, dan Pekerjaan mampu menjelaskan variabel asal. Pengelompokan komponen pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 berdasarkan matriks komponen di rotasi (*Rotated Component Matrix*) yang terbentuk ditunjukkan sebagai berikut.

Tabel 8 *Rotated Component Matrix*

Variabel	Component	
	1	2
RLS	0,691	0,501
TPT	0,671	-0,015
Indeks Literasi Digital	0,467	0,007
TIK	0,904	0,139
Internet	0,892	0,158
Keterampilan Digital	0,052	0,770
Pemberdayaan	0,635	0,479
Pekerjaan	0,056	0,783

Berdasarkan Tabel 8 menunjukkan pengelompokan setiap variabel ke dalam faktor tertentu yang didasarkan nilai terbesar komponen matriks tiap variabel. Dapat dilihat bahwa faktor 1 terdiri dari variabel RLS, TPT, Indeks Literasi Digital, TIK, Internet, dan Pemberdayaan. Faktor 2 terdiri dari variabel Keterampilan Digital dan Pekerjaan. Secara ringkas masing-masing faktor yang terbentuk ditunjukkan sebagai berikut.

Tabel 9 Faktor Baru Data Faktor-Faktor Tingkat Penggunaan *E-Commerce* di Setiap Provinsi Indonesia Tahun 2023

Faktor	Variabel	Keterangan
Pendidikan dan Kualitas Sumber Daya Manusia	X ₁	Rata-rata Lama Sekolah
	X ₂	Tingkat Pengangguran Terbuka
	X ₃	Indeks Literasi Digital
	X ₄	Teknologi Informasi dan Komunikasi
	X ₅	Internet
	X ₇	Pemberdayaan
Keterampilan dan Ketenagakerjaan	X ₆	Keterampilan Digital
	X ₈	Pekerjaan

Berdasarkan Tabel 9 menunjukan bahwa penamaan faktor dilakukan berdasarkan variabel-variabel yang masuk di dalam faktor baru. Untuk faktor pertama yaitu pendidikan dan kualitas sumber daya manusia karena variabel yang masuk didalamnya menjelaskan RLS, TPT, Indeks Literasi Digital, Teknologi Informasi dan Komunikasi, Internet, dan Pemberdayaan. Untuk faktor kedua yaitu Keterampilan dan Ketenagakerjaan karena variabel didalamnya seperti Keterampilan Digital dan Pekerjaan. Komponen transformasi matriks berfungsi untuk membuktikan besarnya nilai korelasi dari faktor yang terbentuk pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023. Berikut komponen transformasi matriks pada data variabel faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 ditunjukkan sebagai berikut.

Tabel 10 Komponen Transformasi Matriks pada Data Faktor-faktor Tingkat Penggunaan *E-Commerce* di Setiap Provinsi Indonesia Tahun 2023

Component	1	2
1	0,890	0,455
2	-0,455	0,890

Berdasarkan Tabel 10 menunjukkan bahwa nilai korelasi pada komponen 1 dan 2 memiliki nilai berturut-turut sebesar 0,890, dimana nilai-nilai tersebut lebih besar dari 0,5 sehingga dapat dikatakan bahwa kedua faktor yang terbentuk dikatakan tepat dalam merangkum delapan variabel yang digunakan pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023.

SIMPULAN

Berdasarkan hasil analisis yang telah dilakukan diperoleh beberapa kesimpulan antara lain sebagai berikut.

1. Semua variabel yang mempengaruhi persentase penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 dapat dikatakan data kurang bervariasi atau homogen.
2. Hasil uji asumsi analisis faktor pada data faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia tahun 2023 memenuhi asumsi distribusi normal secara multivariat, dependen, baik untuk difaktorkan, dan dapat diprediksi serta dianalisis lebih lanjut.
3. Hasil analisis faktor diperoleh 2 komponen utama atau faktor terbentuk yang dapat menjelaskan variabel asal sebesar 61,237%. 2 faktor tersebut yaitu untuk faktor pertama Pendidikan dan Kuliatas Sumber daya manusia karena variabel yang masuk didalamnya menjelaskan Rata-rata Lama Sekolah, Tingkat Pengangguran Terbuka, Indeks Literasi Digital, Teknologi Informasi dan Komunikasi, Internet, dan Pemberdayaan. Sedangkan Faktor kedua atau Keterampilan dan Ketenagakerjaan karena variabel didalamnya seperti Keterampilan Digital dan Pekerjaan.

SARAN

Berdasarkan hasil penelitian, saran kepada peneliti selanjutnya diharapkan agar menambah variabel lain yang berkaitan dengan faktor-faktor yang memengaruhi penggunaan *e-commerce* di setiap provinsi Indonesia termasuk indikator digitalisasi, seperti tingkat literasi digital, akses terhadap infrastruktur internet, dan ketersediaan layanan digital di daerah. Selain itu, pemerintah dapat memberikan kebijakan untuk kelengkapan data penggunaan *e-commerce* di setiap provinsi dengan fokus pada penerapan sistem pelaporan yang terstandar, memanfaatkan teknologi untuk pengumpulan data real-time dan menjalin kerja sama dengan platform *e-commerce*.

REFERENSI

- Badan Pusat Statistik. (2023). Statistik Pendidikan dan Ketahanan Ekonomi Indonesia. *BPS*.
- Bandalos, D. L. (2020). Factor Analysis: Exploratory and Confirmatory. *APA Handbook of Research Methods in Psychology*.
- Beavers, A. S. (2014). Practical Considerations for Using Exploratory Factor Analysis in Educational Research. *Practical Assessment, Research, and Evaluation*, 1-13.
- Bertot, J. C. (2020). *Public Libraries and the Digital Divide: The Role of Public Libraries in the Digital Age*. Library Quarterly.
- Cerny, B. A. (2022). A Study of a Measure of Sampling Adequacy for Factor Analysis. *Journal of Data Analysis*.
- Cohen, J. &. (2020). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. *Routledge*.
- Field, A. (2021). Discovering Statistics Using IBM SPSS Statistics. *SAGE Publications*.
- Floyd, F. J. (2022). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment*, 368-379.
- Hair, J. F. (2019). *Multivariate Data Analysis*. Pearson.
- Hwang, I. S. (2022). Scree Plot Analysis: A Guide to Selecting the Number of Factors. *Statistical Methods in Medical Research*, 802-812.
- Kaiser, H. F. (2020). The General Factor of Personality: A Meta-Analysis. *Journal of Research in Personality*.
- Katadata. (2024). *Wamenkominfo Ungkap Tiga Tantangan Ekonomi Digital di Masa Depan*. Katadata.
- Kementerian Perdagangan RI. (2024). *Kemendag Ramal Transaksi E-Commerce di RI Tembus Rp533 Triliun*. Kementerian Perdagangan RI.
- Koot, H. M. (2019). The Role of Bartlett's Test in Assessing Factor Analysis. *Psychological Methods*, 517-529.
- McKinsey Global Institute. (2021). *The Future of Work: Reskilling and Remote Work in the Digital Age*. McKinsey Global Institute.
- Mishra, P. &. (2022). Multivariate Normality Testing: A New Perspective. *Journal of Statistical Theory and Practice*, 48-63.
- OECD. (2021). *Digital Economy Outlook 2021*. OECD.
- Pew Research Center. (2022). *The Digital Divide Persists Even as Lower-Income Americans Make Gains in Tech Adoption*. Pew Research Center.
- Rummel, R. J. (2021). Applied Factor Analysis. *Random House*.
- Tabachnick, B. G. (2019). *Using Multivariate Statistics*. Pearson.
- World Economic Forum. (2022). *The Global Competitiveness Report 2022*. World Economic Forum.