
Analisis Korelasi Kanonik antara Indikator Sosial Ekonomi dan Indikator Kesehatan serta Gizi di Indonesia Tahun 2020

(Canonical Correlation Analysis between Socioeconomic and Health-Nutrition Indicators in Indonesian in 2020)

Barrafi Qauli Syahnaz¹, Khinaya Rayhanun Jannah¹, Syahma Wati Berutu¹,
Dara Irsalina^{2*}, Ahmad Hadra Zuhri²

¹ Department of Statistics, Universitas Syiah Kuala, Banda Aceh, Indonesia

² Department of Mathematics, Universitas Syiah Kuala, Banda Aceh, Indonesia

E-mail: barrafi@mhs.usk.ac.id; rayhanun@mhs.usk.ac.id; syahma@mhs.usk.ac.id;
dara.irsalina@usk.ac.id*; ahadrazuhri@usk.ac.id;

Abstrak

Ketimpangan dalam indikator kesehatan dan gizi masyarakat di Indonesia masih menjadi isu krusial, terutama pasca pandemi COVID-19 yang memperburuk kondisi sosial-ekonomi. Faktor-faktor seperti pendidikan, ketimpangan pendapatan, kepadatan penduduk, dan daya beli diduga berkontribusi terhadap capaian kesehatan masyarakat. Penelitian ini bertujuan untuk menganalisis hubungan antara indikator sosial-ekonomi dan indikator kesehatan serta gizi di Indonesia tahun 2020. Metode *Canonical Correlation Analysis* (CCA) dipilih karena mampu mengidentifikasi hubungan multivariat antara dua kelompok variabel secara simultan, sehingga cocok untuk mengkaji keterkaitan kompleks antar indikator. Hasil analisis menunjukkan bahwa hanya kombinasi variabel sosial-ekonomi dan kesehatan pada komponen pertama yang menunjukkan hubungan signifikan. Temuan ini menegaskan pentingnya pendekatan kebijakan yang tidak hanya terfokus pada sektor kesehatan, tetapi juga mencakup sektor pendidikan, ekonomi, dan kesejahteraan secara lebih luas.

Abstract

Disparities in health and nutrition indicators among the Indonesian population remain a critical issue, particularly in the aftermath of the COVID-19 pandemic, which has exacerbated socio-economic conditions. Factors such as education, income inequality, population density, and household purchasing power are believed to significantly influence public health outcomes. This study aims to analyze the relationship between socio-economic indicators and health and nutrition indicators in Indonesia in 2020. The Canonical Correlation Analysis (CCA) method was selected for its ability to identify multivariate relationships between two sets of variables simultaneously, making it suitable for examining complex inter-indicator linkages. The analysis revealed that only the first pair of canonical variates—linking socio-economic and health indicators—showed

Informasi Artikel

Sejarah Artikel:

Diajukan 31 Mei 2025

Direvisi 4 Jun 2025

Diterima 18 Jun 2025

Kata Kunci:

Gizi

Indikator sosial-ekonomi

Indonesia

Kesehatan masyarakat

Korelasi Kanonik

Keywords:

Nutrition

Socio-economic indicators

Indonesia

Public health

Canonical correlation

a statistically significant relationship. These findings underscore the importance of policy approaches that go beyond the health sector alone, incorporating education, economic development, and broader social welfare considerations.

1. Pendahuluan

Kesehatan dan gizi masyarakat di Indonesia masih menjadi tantangan penting dalam pembangunan nasional. Ketimpangan akses terhadap layanan kesehatan dan kualitas gizi yang tidak merata menjadi sorotan, terutama pada masa krisis seperti pandemi COVID-19 yang melanda pada tahun 2020. Situasi tersebut memperburuk kondisi sosial-ekonomi dan memperlebar kesenjangan antarwilayah, baik dalam hal kesehatan maupun kesejahteraan. Data dari Badan Pusat Statistik menunjukkan adanya perbedaan signifikan antarwilayah dalam hal angka kematian bayi, harapan hidup, dan proporsi balita bergizi baik, yang mengindikasikan pengaruh kuat dari kondisi sosial-ekonomi masyarakat.

Pendidikan menjadi salah satu faktor penting yang memengaruhi kualitas kesehatan, karena masyarakat yang memiliki tingkat pendidikan lebih tinggi cenderung memiliki pemahaman lebih baik tentang perilaku hidup sehat. Demikian pula, ketimpangan distribusi pendapatan — yang tercermin dalam nilai koefisien Gini berdampak pada akses masyarakat terhadap layanan kesehatan. Selain itu, kepadatan penduduk yang tinggi sering kali tidak diimbangi dengan penyediaan fasilitas kesehatan yang memadai. Kemampuan ekonomi yang tercermin dari pengeluaran per kapita juga turut menentukan sejauh mana masyarakat dapat memenuhi kebutuhan dasar, termasuk pangan bergizi dan layanan kesehatan.

Studi terkait keterkaitan sosial ekonomi dan kesehatan gizi telah banyak dilakukan. Seperti, Safitri dan Indrasari (2009) mempelajari penggunaan *Canonical Correlation Analysis* CCA dalam mengkaji hubungan antara karakteristik sosial-ekonomi dan perilaku kesehatan di Kota Pati— dengan sembilan indikator masing-masing dari kedua kelompok variabel, Sari, et. al (2018) menerapkan CCA untuk mengkaji keterkaitan antara dimensi pendidikan, kesehatan, dan kualitas hidup rumah tangga dengan derajat kemiskinan di Indonesia, dengan koefisien korelasi kanonik sebesar 0,810. Namun, penelitian-penelitian tersebut terbatas pada satu daerah saja. Pada penelitian ini, penulis tertarik untuk melihat bagaimana keterkaitan multivariat antara indikator sosial ekonomi dan kesehatan gizi menggunakan CCA di Indonesia.

Adapun tujuan dari penelitian ini adalah untuk mengkaji hubungan antara rata-rata lama sekolah, koefisien Gini, kepadatan penduduk, dan pengeluaran per kapita dengan angka kematian bayi, umur harapan hidup, persalinan yang ditangani tenaga medis, dan status gizi balita di Indonesia pada tahun 2020. Diharapkan hasil dari penelitian ini dapat menjadi landasan dalam perumusan kebijakan kesehatan dan sosial yang lebih tepat sasaran, serta memberikan pemahaman baru bagi para pengambil kebijakan, akademisi, dan masyarakat umum tentang pentingnya integrasi kebijakan sosial dan kesehatan dalam pembangunan nasional.

2. Tinjauan Pustaka

2.1. Statistika Deskriptif

Statistik deskriptif digunakan untuk menggambarkan atau meringkas data yang diperoleh dari suatu penelitian. Melalui ukuran seperti rata-rata, median, modus, standar deviasi, dan range, peneliti dapat memperoleh gambaran umum mengenai karakteristik data sebelum dilakukan

analisis lebih lanjut [1]. Statistik deskriptif juga penting untuk mendeteksi nilai-nilai ekstrem dan memberikan indikasi awal distribusi data.

2.2. Uji Normalitas

Uji normalitas merupakan prosedur statistik yang digunakan untuk menentukan apakah data dalam suatu variabel atau kumpulan variabel mengikuti distribusi normal. Distribusi normal merupakan salah satu asumsi penting dalam banyak analisis statistik parametrik seperti regresi, ANOVA, dan MANOVA. Oleh karena itu, sebelum melakukan analisis lanjutan, penting untuk memastikan bahwa data memiliki distribusi yang mendekati normal [2].

2.2.1. Uji Normalitas Univariat

Analisis Multivariat merupakan suatu analisis yang digunakan apabila pengukuran setiap variabel dilakukan secara bersamaan atau serempak [3]. Uji normalitas univariat digunakan untuk menguji apakah distribusi suatu variabel tunggal mengikuti distribusi normal. Uji ini penting karena banyak teknik statistik parametrik mensyaratkan data berdistribusi normal secara individual sebelum analisis dilakukan [4].

2.2.2. Uji Normalitas Multivariat

Uji normalitas multivariat digunakan untuk mengetahui apakah sejumlah variabel yang dianalisis secara bersama-sama memiliki distribusi normal multivariat. Uji ini sangat penting dalam analisis multivariat seperti MANOVA, analisis diskriminan, atau analisis faktor, karena asumsi normalitas multivariat memengaruhi validitas hasil uji statistik tersebut [5].

2.3. Uji Multikolinearitas

Multikolinearitas adalah suatu kondisi di mana terjadi korelasi yang tinggi antar variabel independen dalam suatu model regresi. Keberadaan multikolinearitas dapat menyebabkan ketidakstabilan dalam estimasi parameter regresi, meningkatkan standar error, dan menyulitkan interpretasi pengaruh masing-masing variabel independen. Secara umum, hubungan antar variabel dapat diidentifikasi dengan beberapa metode statistika, seperti Biplot [6], Khi-Kuadrat [7], [8] dan korelasi. Untuk mendeteksi multikolinearitas pada penelitian ini, dua indikator utama digunakan yaitu *Variance Inflation Factor* (VIF) dan *Tolerance*. Nilai VIF yang lebih dari 10 atau tolerance kurang dari 0,1 umumnya dianggap sebagai indikasi adanya multikolinearitas tinggi [9]. Jika multikolinearitas ditemukan, maka solusi yang umum dilakukan adalah menghapus salah satu variabel, melakukan transformasi, atau menggunakan metode analisis alternatif seperti regresi ridge.

2.4. Eigen values and canonical correlation

Dalam analisis statistik multivariat, eigen values (nilai eigen) menunjukkan jumlah varians yang dapat dijelaskan oleh masing-masing komponen atau fungsi dalam suatu analisis, seperti Principal Component Analysis (PCA) atau Canonical Correlation Analysis (CCA). Nilai eigen digunakan untuk menentukan seberapa penting suatu dimensi atau fungsi dalam menjelaskan hubungan antar variabel [10].

Eigen value yang besar menunjukkan bahwa dimensi atau fungsi tersebut menjelaskan proporsi besar dari total varians data. Dalam konteks CCA, nilai eigen terkait langsung dengan kekuatan hubungan antar dua set variabel. CCA adalah teknik statistik yang digunakan untuk

mengukur dan menganalisis hubungan antara dua set variabel multivariat. Tujuan utama dari CCA adalah untuk menemukan pasangan kombinasi linier dari dua set variabel (misalnya, X dan Y) yang memiliki korelasi maksimal satu sama lain.

2.5. Dimension reduction analysis (Uji Individu)

Dimension Reduction Analysis adalah teknik statistik yang digunakan untuk menyederhanakan struktur data multivariat dengan cara mengurangi jumlah variabel independen (dimensi) tanpa menghilangkan informasi penting yang terkandung dalam data tersebut [11]. Tujuan utamanya adalah untuk mengidentifikasi dan mengekstraksi struktur laten atau faktor utama yang mendasari hubungan antar variabel. Dalam konteks CCA, uji individu pada dimension reduction mengacu pada pengujian signifikansi dari masing-masing canonical function (pasangan variabel kanonik). Uji ini dilakukan untuk menentukan berapa banyak dimensi (canonical functions) yang signifikan dalam menjelaskan hubungan antara dua set variabel. Setiap dimensi diuji secara individual untuk melihat apakah hubungan kanonik tersebut berbeda secara signifikan dari nol.

2.6. Multivariate test of significance (Uji Kelompok)

Uji signifikansi multivariat atau Multivariate Test of Significance merupakan langkah penting dalam analisis multivariat yang bertujuan untuk mengevaluasi apakah secara keseluruhan terdapat hubungan atau perbedaan yang signifikan antara dua atau lebih kelompok variabel. Dalam konteks seperti CCA atau MANOVA, uji ini digunakan untuk menilai apakah set variabel independen secara bersama-sama berhubungan secara signifikan dengan set variabel dependen [12].

2.7. Analisis Korelasi Kanonik (CCA)

CCA merupakan teknik statistik multivariat yang digunakan untuk mengkaji hubungan antara dua set variabel yang berbeda, yaitu variabel dependen ($Y_1, Y_2, \dots, Y_n$) dan variabel independen ($X_1, X_2, \dots, X_n$). Metode ini pertama kali diperkenalkan oleh Hotelling (1936) dan memungkinkan peneliti untuk mengevaluasi korelasi antara kombinasi linier dari dua kelompok variabel. Dalam konteks penelitian ini, CCA digunakan untuk menganalisis sejauh mana indikator sosial-ekonomi secara simultan berkaitan dengan indikator kesehatan dan gizi. Metode ini sangat relevan dalam studi interdisipliner yang melibatkan banyak dimensi, seperti hubungan antara pembangunan sosial dan kesehatan publik.

Asumsi yang digunakan dalam CCA:

1. Hubungan antara variabel dependen dengan variabel independen bersifat linier.
2. Korelasi kanonik adalah hubungan linier antar variabel kanonik
3. Variabel dependen dan independen berdistribusi Normal Multivariat

3. Metode Penelitian

Adapun tahapan penelitian dalam kajian ini meliputi:

1. Pengumpulan data sosial ekonomi dan kesehatan dari 34 provinsi di Indonesia. Penelitian ini menggunakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) Pusat Tahun 2020. Data tersebut mencakup indikator sosial ekonomi dan kesehatan dari 34 provinsi di Indonesia, yang diperoleh melalui publikasi resmi lembaga terkait seperti Kementerian Kesehatan dan BPS sendiri, dengan akses berbasis internet.

2. Menentukan variabel independen dan dependen. Pada penelitian ini variabel independen merupakan sosial ekonomi dan variabel dependen merupakan kesehatan serta gizi.
3. Menentukan indikator variabel dependen dan independen yang akan digunakan. Indikator dari variabel dependen meliputi Angka Kematian Bayi (AKB), Usia Harapan Hidup (UHH), Persentase Persalinan yang Ditangani oleh Dokter/Bidan, dan Status Gizi Baik pada Balita. Sementara itu, indikator dari variabel independen terdiri dari Rata-Rata Lama Sekolah (RLS), Koefisien Gini, Kepadatan Penduduk per Kilometer Persegi, dan Pengeluaran Rata-Rata Per Kapita Sebulan (dalam Rupiah).
4. Melakukan analisis CCA dan jumlah pasangan fungsi kanonik. Analisis CCA digunakan untuk mengidentifikasi dan mengukur hubungan linier antara dua set variabel. Sedangkan jumlah fungsi kanonik digunakan untuk melihat korelasi terkuat.
5. Interpretasi hasil. Hasil analisis yang diperoleh akan diinterpretasikan menggunakan koefisien kanonik, muatan kanonik (canonical loadings), dan uji signifikansi (seperti uji Wilks' Lambda) untuk menentukan kekuatan dan relevansi statistik dari hubungan yang ditemukan.

4. Analisis Pembahasan

4.1. Analisis Deskriptif Data Provinsi

Tabel 1 Deskriptif data provinsi

Variabel	Min	Mean	Median	Max
Y1	103,8000	19,7400	17,2300	38,1700
Y2	61,2200	72,2600	72,6100	81,9200
Y3	68,4900	96,5100	96,1100	100,0000
Y4	66,0000	73,5700	72,9500	84,4000
X1	6,9600	9,0800	9,1500	11,1700
X2	0,2900	0,3400	0,3300	0,4300
X3	9,0000	739,2400	102,5000	15907,0000
X4	1032,8900	1273,2100	1226,7200	1992,8500
Provinsi	Object			

Tabel 1 menyajikan statistik deskriptif untuk berbagai peubah yang diukur di sejumlah provinsi. Peubah Y1 memiliki nilai minimum sebesar 10,38 (tercatat pada DKI Jakarta) dan maksimum 38,17 (tercatat pada Papua), dengan nilai rata-rata 19,74 dan median 17,23. Peubah Y2 menunjukkan rentang nilai antara 61,22 hingga 81,92, dengan nilai rata-rata 72,26 dan median 72,61. Peubah Y3 memiliki nilai minimum 68,49 dan maksimum 100,00, dengan rata-rata 96,51 dan median 96,11. Sementara itu, peubah Y4 menunjukkan rentang nilai antara 66,00 dan 84,40, dengan rata-rata 73,57 dan median 72,95. Untuk peubah X1, nilai minimum tercatat sebesar 6,96 dan maksimum 11,17, dengan rata-rata 9,08 dan median 9,15. Peubah X2 menunjukkan rentang nilai yang lebih sempit, yaitu antara 0,29 hingga 0,43, dengan rata-rata 0,34. Peubah X3 memiliki distribusi nilai yang sangat lebar, dengan nilai minimum 9,00 dan maksimum 15907,00, rata-rata 739,24, serta median 102,50. Terakhir, peubah X4 memiliki nilai minimum 1032,89 dan maksimum 1992,85, dengan rata-rata 1273,21 dan median 1226,72.

4.2. Uji Normalitas

4.2.1. Uji Normalitas Multivariat terhadap X

Tabel 2 Hasil uji normalitas multivariat

<i>Test</i>	<i>H</i>	<i>p-value</i>	<i>MVN</i>
Royston	21,3544	$3,78 \times 10^{-5}$	NO

Berdasarkan hasil uji normalitas multivariat yang disajikan pada Tabel 2, didapatkan nilai statistik uji Royston sebesar 21,35443 dengan p-value sebesar $3,78 \times 10^{-5}$. Nilai p-value yang sangat kecil (jauh lebih kecil dari 0,05) menunjukkan bahwa hipotesis nol (H_0) yang menyatakan bahwa data berdistribusi normal multivariat ditolak. Oleh karena itu, dapat disimpulkan bahwa data yang dianalisis tidak berdistribusi normal multivariat. Kegagalan memenuhi asumsi normalitas multivariat ini memiliki implikasi penting terhadap analisis korelasi kanonik, karena asumsi normalitas merupakan prasyarat utama untuk validitas uji statistik dan interpretasi hasil. Pelanggaran asumsi ini dapat menyebabkan estimasi koefisien kanonik dan nilai signifikansi menjadi kurang akurat atau bias.

Tabel 3 Hasil uji normalitas univariat

<i>Test</i>	<i>Variable</i>	<i>Statistic</i>	<i>p-value</i>	<i>Normality</i>
Anderson-Darling	Y1	1,6389	0,0003	NO
Anderson-Darling	Y2	0,7372	0,0495	NO
Anderson-Darling	Y3	2,6673	<0,0010	NO
Anderson-Darling	Y4	0,4455	0,2668	YES

Berdasarkan hasil uji normalitas univariat pada Tabel 3, variabel X_1 , X_2 , dan X_3 tidak berdistribusi normal univariat, karena p-value untuk masing-masing variabel (X_1 : 0,0003, X_2 : 0,0495, X_3 : <0,0010) lebih kecil dari α (0,05). Sementara itu, variabel X_4 berdistribusi normal univariat, karena p-value (0,2668) lebih besar dari α (0,05).

4.2.2. Uji Normalitas Multivariat terhadap Y

Tabel 4 Hasil uji normalitas multivariat

<i>Test</i>	<i>H</i>	<i>p-value</i>	<i>MVN</i>
Royston	73,3156	$4,99 \times 10^{-15}$	NO

Hasil uji normalitas multivariat yang disajikan dalam Tabel 4 menggunakan uji Royston menunjukkan nilai H sebesar 73,31567 dengan p-value sebesar $4,99 \times 10^{-15}$. Berdasarkan p-value yang sangat kecil (jauh lebih kecil dari level signifikansi, sebesar 0,05), dapat disimpulkan bahwa hipotesis nol (yang menyatakan bahwa data mengikuti distribusi normal multivariat) ditolak. Oleh karena itu, hasil uji ini menunjukkan bahwa data tidak mengikuti distribusi normal multivariat (MVN = NO).

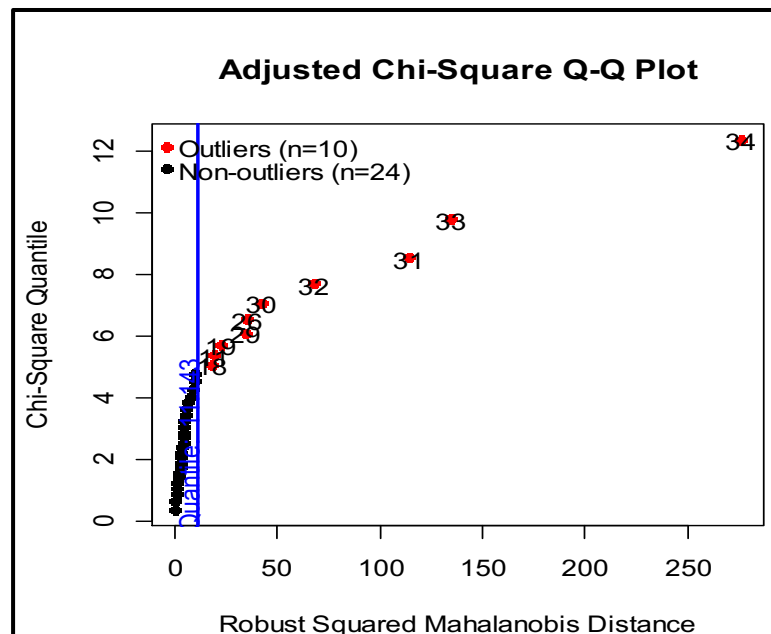
Tabel 5 Hasil uji normalitas univariat

<i>Test</i>	<i>Variable</i>	<i>Statistic</i>	<i>p-value</i>	<i>Normality</i>
Anderson-Darling	X1	0,1908	0,8910	YES
Anderson-Darling	X2	1,1431	0,0047	NO
Anderson-Darling	X3	9,7100	<0,0010	NO
Anderson-Darling	X4	2,2962	<0,0010	NO

Hasil uji normalitas univariat yang disajikan dalam Tabel 5 menggunakan uji Anderson-Darling menunjukkan bahwa untuk variabel X_1 , nilai statistik sebesar 0,1908 dengan p -value 0,891, yang lebih besar dari 0,05, mengindikasikan bahwa data untuk variabel X_1 berdistribusi normal univariat (Normality = YES). Namun, untuk variabel X_2 , X_3 , dan X_4 , p -value yang sangat kecil (X_2 : 0,0047, X_3 : <0,0010, X_4 : <0,0010), jauh lebih kecil dari 0,05, menunjukkan bahwa hipotesis nol ditolak, yang berarti data untuk ketiga variabel tersebut tidak berdistribusi normal univariat (Normality = NO). Dengan demikian, hanya variabel X_1 yang memenuhi asumsi normalitas univariat, sementara variabel X_2 , X_3 , dan X_4 tidak berdistribusi normal.

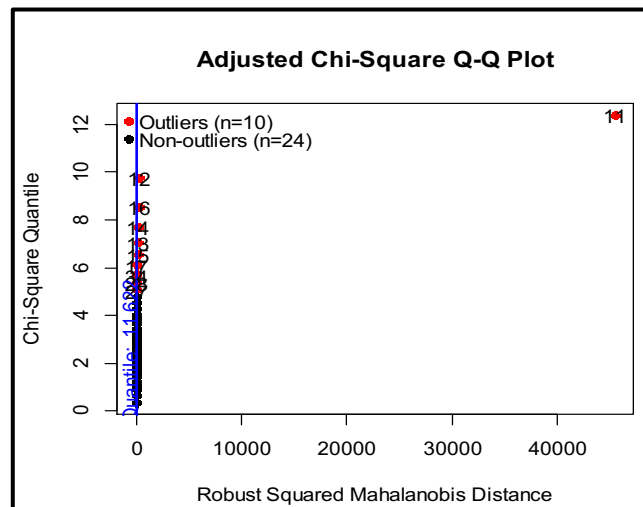
4.3. Melihat Outlier

Outlier adalah observasi atau data yang memiliki nilai ekstrem dan berbeda secara signifikan dari observasi lainnya. Dalam konteks analisis multivariat, keberadaan outlier dapat sangat memengaruhi hasil analisis karena outlier dapat menyebabkan distorsi dalam estimasi parameter dan struktur hubungan antar variabel [13].



Gambar 1 Q-Q Plot variabel X.

Berdasarkan gambar 1 dapat dilihat bahwa variabel X (RLS, Koefisien Gini, Kepadatan Penduduk Per-KM, dan Pengeluaran Rata-rata Perkapita Sebulan) terdiri dari 34 data dengan 10 nilai outliers dan 24 non-outliers.



Gambar 2 Q-Q Plot variabel Y.

Berdasarkan Gambar 2 dapat dilihat bahwa variabel (Angka Kematian Bayi, UHH, Persalinan Ditolong Dokter/Bidan, Balita Berstatus Gizi Baik) terdiri dari 34 data dengan 10 nilai outliers dan 24 non-outliers.

4.4. Uji Multikolinearitas

Tabel 6 Hasil uji korelasi antar variabel

	<i>Y1</i>	<i>Y2</i>	<i>Y3</i>	<i>Y4</i>	<i>X1</i>	<i>X2</i>	<i>X3</i>	<i>X4</i>
<i>Y1</i>	1	-0,8155	-0,7218	-0,5688	-0,3002	-0,0356	-0,3031	0,0714
<i>Y2</i>	-0,8155	1	0,6807	0,5359	0,6948	0,1934	0,4981	0,1047
<i>Y3</i>	-0,7218	0,6807	1	0,3044	0,2384	-0,0412	0,1794	-0,1913
<i>Y4</i>	-0,5688	0,5359	0,3044	1	0,2410	0,2628	0,2946	0,2660
<i>X1</i>	-0,3002	0,6948	0,2384	0,2410	1	-0,0126	0,4380	0,3290
<i>X2</i>	-0,0356	0,1934	-0,0412	0,2628	-0,0126	1	0,4603	0,2052
<i>X3</i>	-0,3031	0,4981	0,1794	0,2946	0,4380	0,4603	1	0,5021
<i>X4</i>	0,0714	0,1047	-0,1913	0,2660	0,3290	0,2052	0,5021	1

Berdasarkan hasil korelasi antar variabel yang disajikan dalam Tabel 6, terdapat beberapa pasangan variabel dengan korelasi lebih besar dari 0,5 yang menunjukkan adanya multikolinearitas antar variabel. Korelasi kuat ditemukan antara variabel *Y1* dan *Y2* (-0,8155), *Y2* dan *X1* (0,6948), *Y2* dan *Y3* (0,6807), serta *X1* dan *X2* (0,6948). Korelasi yang lebih besar dari 0,5 ini mengindikasikan adanya hubungan yang kuat antara variabel-variabel tersebut, yang dapat mempengaruhi model statistik jika tidak ditangani dengan baik, seperti dalam analisis regresi. Multikolinearitas yang tinggi dapat menyebabkan estimasi koefisien yang tidak stabil dan interpretasi model yang kurang akurat.

4.5. Nilai Eigen dan Korelasi Kanonik

Tabel 7 Hasil koefisien korelasi kanonik yang terstandarisasi (Rho)

Komponen	Korelasi Kanonik (Rho)
1	0,9077
2	0,5440
3	0,1840
4	0,0540

Berdasarkan output yang disajikan pada Tabel 7, dapat diketahui bahwa komponen pertama memiliki nilai korelasi kanonik sebesar 0,9077, yang menunjukkan adanya hubungan linier yang sangat kuat antara dua set variabel pada komponen pertama. Komponen kedua memiliki nilai korelasi kanonik sebesar 0,5440, yang menunjukkan adanya hubungan linier yang sedang antara dua set variabel pada komponen kedua. Komponen ketiga memiliki nilai korelasi kanonik sebesar 0,1840, yang menunjukkan hubungan linier yang lemah antara kedua set variabel pada komponen ketiga. Terakhir, komponen keempat memiliki nilai korelasi kanonik sebesar 0,0540, yang menunjukkan hubungan linier yang sangat lemah antara kedua set variabel pada komponen keempat. Secara keseluruhan, hasil ini menunjukkan bahwa hubungan antar set variabel semakin lemah pada komponen yang lebih tinggi.

4.6. Dimension reduction analysis (Uji Individu)

Tabel 8 Hasil uji signifikansi korelasi kanonik secara individu

Komponen	Statistic	Approx	df1	df2	p-value
1 to 4	0,1195	5,0279	16	80,07	$4,86 \times 10^{-7}$
2 to 4	0,6782	1,2658	9	65,86	0,2722
3 to 4	0,9633	0,2641	4	56	0,9000
4 to 4	0,9971	0,0849	1	29	0,7728

Berdasarkan hasil uji signifikansi korelasi kanonik yang disajikan dalam Tabel 8, dapat disimpulkan bahwa korelasi kanonik antara komponen pertama dan keempat (1 to 4) signifikan secara individu dengan p-value sebesar $4,86 \times 10^{-7}$ yang lebih kecil dari α (0,05), sehingga H_0 ditolak. Sebaliknya, korelasi kanonik antara komponen kedua dan keempat (2 to 4), ketiga dan keempat (3 to 4), serta keempat dengan dirinya sendiri (4 to 4) tidak signifikan secara individu karena p-value masing-masing lebih besar dari α (0,05), yaitu 0,2722, 0,9000, dan 0,7728, sehingga H_0 diterima pada ketiga pasangan tersebut.

Dengan ditemukannya hubungan yang signifikan secara statistik antara komponen pertama dari indikator sosial ekonomi dan komponen keempat dari indikator kesehatan dan gizi, dapat diinterpretasikan bahwa terdapat keterkaitan yang kuat antara kombinasi variabel-variabel dalam komponen tersebut. Artinya, variabel-variabel sosial ekonomi yang memiliki kontribusi dominan pada komponen pertama berpengaruh secara nyata terhadap indikator-indikator kesehatan dan gizi yang tergabung dalam komponen keempat.

Dengan demikian, perubahan pada kondisi sosial ekonomi tersebut secara signifikan dapat mempengaruhi kondisi kesehatan dan gizi masyarakat. Misalnya, perbaikan dalam tingkat pendidikan dan pendapatan dapat berkontribusi pada penurunan angka stunting atau peningkatan status gizi. Sebaliknya, komponen-komponen lainnya tidak menunjukkan hubungan yang

signifikan, yang berarti bahwa kombinasi variabel pada komponen kedua dan ketiga tidak memiliki pengaruh yang cukup kuat terhadap indikator pada komponen keempat.

4.7. Uji Signifikansi Multivariat (Uji Kelompok)

Tabel 9 Hasil uji signifikansi korelasi kanonik secara serempak

Test	Df	Statistic	Approx F	num Df	den Df	p-value
<i>Pillai</i>	4	1,1566	2,9491	16	116	$4,108 \times 10^{-4}$ ***
<i>Wilks</i>	4	0,1195	5,0279	16	80,069	$4,858 \times 10^{-7}$ ***
<i>Roy</i>	4	4,6774	33,9110	4	29	$1,503 \times 10^{-10}$ ***
<i>Hotelling-Lawley</i>	4	-	7,8641	16	98	$1,225 \times 10^{-11}$ ***

Berdasarkan hasil uji signifikansi korelasi kanonik secara serempak yang disajikan dalam Tabel 9, dapat disimpulkan bahwa dengan signifikansi $\alpha = 0,01$, korelasi kanonik signifikan secara bersama-sama pada seluruh uji yang dilakukan. Uji Pillai menghasilkan p-value sebesar $4,108 \times 10^{-4}$, Uji Wilks menghasilkan p-value $4,858 \times 10^{-7}$, Uji Roy menghasilkan p-value $1,503 \times 10^{-10}$, dan Uji Hotelling-Lawley menghasilkan p-value $1,225 \times 10^{-11}$, yang semuanya lebih kecil dari α (0,05). Oleh karena itu, H_0 ditolak pada semua uji tersebut, yang menunjukkan bahwa korelasi kanonik antar variabel dalam model ini signifikan secara serempak.

5. Kesimpulan

Berdasarkan hasil analisis, maka kesimpulan yang diperoleh adalah sebagai berikut.

1. Terdapat hubungan yang signifikan pada signifikansi $\alpha = 0.01$ indikator sosial ekonomi dan indikator kesehatan serta gizi di Indonesia. Temuan ini menunjukkan bahwa variabel sosial-ekonomi secara simultan memiliki pengaruh yang bermakna terhadap kondisi kesehatan dan gizi masyarakat di berbagai provinsi di Indonesia.
2. Komponen pertama dalam analisis korelasi kanonik menunjukkan korelasi paling kuat dengan nilai 0,9077, yang mengindikasikan adanya hubungan linier yang sangat tinggi antara kedua kelompok indikator.
3. Dari hasil uji signifikansi individu, hanya komponen pertama yang terbukti signifikan, sedangkan komponen kedua, ketiga, dan keempat tidak menunjukkan hubungan yang signifikan secara statistik.
4. Berdasarkan hasil analisis, indikator rata-rata lama sekolah (RLS) dan koefisien Gini merupakan dua variabel sosial-ekonomi yang paling berkontribusi dalam fungsi kanonik pertama. Kedua indikator ini berasosiasi kuat dengan indikator kesehatan seperti angka kematian bayi dan persentase balita dengan status gizi baik, yang menandakan bahwa tingkat pendidikan dan ketimpangan pendapatan memegang peranan penting dalam menentukan kualitas kesehatan suatu wilayah.
5. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa untuk mencapai peningkatan kondisi kesehatan dan gizi masyarakat, diperlukan pendekatan kebijakan yang tidak hanya fokus pada sektor kesehatan semata, tetapi juga mengintegrasikan upaya di sektor pendidikan, pengentasan ketimpangan ekonomi, dan peningkatan kesejahteraan masyarakat secara menyeluruh. Pendekatan terpadu ini diharapkan mampu menciptakan pembangunan kesehatan yang lebih inklusif dan berkeadilan di seluruh wilayah Indonesia.

Daftar Pustaka

- [1] M. Tarigan and D. F. Silaban, *Statistika Deskriptif*, JINTAN *J. Ilmu Keperawatan*, vol. 4, no. 2, pp. 187–195, Jul. 2024.
- [2] I. Sintia, M. D. Pasarella, and D. A. Nohe, 'Perbandingan Tingkat Konsistensi Uji Distribusi Normalitas pada Kasus Tingkat Pengangguran di Jawa', *Pros. Semin. Nas. Mat. dan Stat.*, vol. 2, May. 2022.
- [3] M. S. Yana, L. Setiawan, E. M. Ulfa, A. Rusyana. 'Penerapan Metode K-Means dalam Pengelompokan Wilayah Menurut Intensitas Kejadian Bencana Alam di Indonesia Tahun 2013-2018', *Journal of Data Analysis*, vol. 1, no 2, pp. 93-102, Dec. 2018.
- [4] B. N. Suryawati, L. Wardani, I. Kusmayadi, and S. Sarmo, *Analisis Varians Multivariat terhadap Return dan Risiko Portofolio yang Ditentukan dengan Elton Grueber Padberg Model pada Empat Kelompok Indeks (LQ 45, SRI-KEHATI, JII, DAN ISSI)*.
- [5] N. Bakhary, K. Yahya, and C. N. Ng, 'Univariate Artificial Neural Network in Forecasting Demand of Low Cost House in Petaling Jaya', *J. Teknol. Sci. Eng.*, pp. 1–16, 2004.
- [6] A. Rusyana, N. Nurhasanah, H. Hizir, 'Identification of Non-Oil and Gas Main Commodity Exports in Indonesia to Major Destination Countries Using PCA Biplot and CVA Biplot', *Journal of Data Analysis*, vol. 7, no 2, pp.65-80, Dec. 2024.
- [7] Z. M. Kesuma, S. Rusdiana, A. Rusyana, L. R. Siregar, R. Rosadi. 'Aplikasi Analisis Korespondensi Berganda Terhadap Status Gizi Remaja di Kota Banda Aceh', *Buletin Penelitian Kesehatan*, vol. 47, no 1, pp:47-54, Jun. 2019.
- [8] M. Mahmudi, M. Fahmi, N. Elfurqany, S. Sarah, A. Rusyana, 'Analisis korespondensi pada korban kecelakaan lalu lintas berdasarkan provinsi di Indonesia', *JMPM: Jurnal Matematika dan Pendidikan Matematika*, vol. 4, no 1, pp:49-60, Mar. 2019.
- [9] S. Lestari, F. Oktavani, A. Wijaya, S. Anwar, 'Hubungan Faktor Multidimensi Terhadap Derajat Kemiskinan di Indonesia Dengan Analisis Korelasi Kanonik', *Journal of Data Analysis*, Agustus. 2021.
- [10] J. Frost, *Multicollinearity in Regression Analysis: Problems, Detection, and Solutions*, Statistics By Jim. Accessed: May 14, 2025.
- [11] J. M. Rickman, Y. Wang, A. D. Rollett, M. P. Harmer, and C. Compson, 'Data analytics using canonical correlation analysis and Monte Carlo simulation', *Npj Comput. Mater.*, vol. 3, no. 1, pp. 1–6, Jul. 2017.
- [12] S. Nanga *et al.*, 'Review of Dimension Reduction Methods', *J. Data Anal. Inf. Process.*, vol. 09, no. 03, pp. 189–231, 2021, doi: 10.4236/jdaip.2021.93013.
- [13] L. Budiarti and B. Warsito, *Analisis Intervensi dan Deteksi Outlier pada Data Wisatawan Domestik (Studi Kasus di Daerah Istimewa Yogyakarta)*.