

PENGARUH INDIKATOR SOSIAL EKONOMI TERHADAP KEPEMILIKAN RUMAH DI INDONESIA: PENDEKATAN *PRINCIPAL COMPONENT ANALYSIS* (PCA) DAN REGRESI POLINOMIAL

Dina Asia Putri^{1*}, Ika Nur Laily Fitriana²

^{1,2} Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas Terbuka, Kota Tangerang Selatan, Indonesia

*Penulis korespondensi: dinaasiap@gmail.com

ABSTRAK

Kepemilikan rumah merupakan isu yang bersifat multidimensi yang dipengaruhi oleh berbagai indikator, baik dari aspek sosial dan ekonomi, serta faktor internal dan eksternal lainnya. Penelitian ini bertujuan untuk menganalisis pengaruh indikator sosial ekonomi terhadap persentase kepemilikan rumah di Indonesia pada tahun 2023 menggunakan pendekatan *Principal Component Analysis* (PCA) dan regresi polinomial. PCA digunakan untuk mereduksi lima variabel sosial ekonomi menjadi dua komponen utama dan mengatasi permasalahan multikolinearitas. Komponen hasil PCA kemudian digunakan dalam model regresi polinomial orde dua, dengan memasukkan komponen linear dan kuadratnya. Hasil analisis menunjukkan bahwa komponen pertama (FAC1), yang merepresentasikan dimensi kesejahteraan ekonomi (tingginya laju PDRB dan persentase kemiskinan serta rendahnya persentase pengangguran) berpengaruh positif dan signifikan ($p < 0,05$) terhadap kepemilikan rumah dengan pola hubungan nonlinier. Sebaliknya, komponen kedua (FAC2), yang berkaitan dengan kepadatan penduduk dan *gini ratio*, tidak menunjukkan pengaruh yang signifikan, karenanya dikeluarkan dari model final dan dilakukan iterasi kembali. Model regresi polinomial final yang melibatkan FAC1 dan bentuk kuadratnya menghasilkan nilai R^2 sebesar 0,561, menunjukkan bahwa model mampu menjelaskan lebih dari separuh variabilitas kepemilikan rumah. Temuan ini mengindikasikan bahwa peningkatan kesejahteraan ekonomi berdampak pada peningkatan kepemilikan rumah, namun efeknya tidak selalu linear dan dapat menurun setelah titik tertentu. Di sisi lain, kemiskinan juga menunjukkan hubungan positif terhadap kepemilikan rumah, yang menjadi temuan menarik dan memerlukan kajian lebih lanjut untuk memahami dinamika sosial ekonomi yang mendasarinya.

Kata kunci: kepemilikan rumah, indikator sosial ekonomi, regresi polinomial, *principal component analysis*, multikolinearitas.

1 PENDAHULUAN

Dalam teori hierarki kebutuhan Maslow, tempat tinggal merupakan salah satu kebutuhan dasar manusia karena berhubungan langsung dengan rasa aman dan stabilitas hidup seseorang maupun keluarganya (Kim & Kim, 2017). Rumah bukan hanya sarana kehidupan semata, tetapi lebih merupakan proses bermukim, yaitu kehadiran manusia sebagai penghuni dalam menciptakan ruang hidup dalam rumah dan lingkungan sekitarnya (Yustie et al., 2024). Namun pada kenyataannya, tidak semua masyarakat mampu memenuhi kebutuhan tempat tinggal karena berbagai kendala yang dihadapi, baik sosial, ekonomi, dan demografi. Sebagian diantaranya lebih memilih untuk menyewa tempat tinggal (Mire, 2024).

Harga rumah di Indonesia terus meningkat dari tahun ke tahun, namun kenaikan ini tidak selalu diimbangi dengan daya beli masyarakat. Survei *Bank Indonesia Residential Property Price* pada 2022 menunjukkan bahwa Indeks Harga Properti Residensial (IHPR) naik menjadi 2,00% (yoy) meskipun penjualan mengalami perlambatan (Bank Indonesia, 2023).

Kesenjangan antara kenaikan harga properti dan pendapatan menyebabkan daya beli masyarakat semakin tertekan, terutama bagi masyarakat menengah ke bawah. Menurut data Badan Pusat Statistik (BPS) tahun 2020, rata-rata Upah Minimum Regional (UMR) di Indonesia berada pada angka Rp2.672.371. Di sisi lain, menurut data dari situs kpr.online yang dihimpun CIMB Niaga, harga rumah subsidi di Indonesia berada pada kisaran Rp150,5 juta hingga Rp219 juta tergantung pada provinsi (CIMB Niaga, 2021). Perbedaan yang cukup besar antara pendapatan dan harga rumah, bahkan untuk rumah subsidi, menunjukkan bahwa kepemilikan rumah sulit dijangkau oleh masyarakat dengan penghasilan rendah (Satoto, 2023). Penyediaan perumahan pun tidak hanya sebatas membangun unit hunian, tetapi juga harus memperhatikan aspek kemanusiaan, kelayakhunian, keamanan, dan kenyamanan. Perumahan ideal juga diharapkan mampu memenuhi standar kesehatan lingkungan serta mendukung keberlangsungan sosial, ekonomi, dan budaya masyarakat (Natalia & Wunas, 2016).

Kepemilikan rumah tidak hanya dipengaruhi oleh kesenjangan antara pendapatan dan harga, tetapi juga oleh tantangan struktural yang bersifat sosial ekonomi. Salah satunya, faktor kemiskinan akan menghambat kemampuan menabung dan membayar uang muka untuk membeli rumah, serta membayar cicilan hipotek yang lebih tinggi (Hafidzi et al., 2024). Selain kemiskinan, faktor sosial ekonomi lainnya juga berpotensi memengaruhi kepemilikan rumah. Misalnya, sebuah penelitian menunjukkan bahwa kepadatan penduduk berpengaruh negatif dan signifikan terhadap pertumbuhan ekonomi di Samarinda artinya jika kepadatan penduduk meningkat maka akan menurun pertumbuhan ekonomi di Samarinda (Yunianto, 2021). Hal ini menekankan perlunya analisis yang lebih lanjut pengaruh indikator sosial ekonomi lain seperti persentase pengangguran terbuka, kemiskinan, kepadatan penduduk, laju PDRB per kapita, gini ratio terhadap kepemilikan rumah.

Dengan mempertimbangkan kemungkinan adanya hubungan yang tidak bersifat linier antar variabel tersebut, data observasi yang tidak berdistribusi normal, dan tingkat multikolinearitas yang cenderung tinggi pada data sosial-ekonomi, maka diperlukan pendekatan statistik yang mampu menangkap pola hubungan yang lebih kompleks dan memberikan interpretasi yang lebih baik. Salah satu metode yang dapat digunakan untuk tujuan tersebut adalah regresi polinomial karena mampu menggambarkan pola hubungan nonlinier antara variabel independen dan variabel dependen (Eka et al., 2021). Untuk mengatasi permasalahan multikolinearitas dan menyederhanakan kompleksitas data sosial ekonomi, penelitian ini juga menerapkan *Principal Component Analysis (PCA)* sebagai tahap awal untuk mereduksi dimensi dan mengekstrak faktor utama dari indikator-indikator sosial ekonomi yang saling berkorelasi (Ramadhan et al., 2024). Selain itu, pendekatan PCA (*Principal Component Analysis*) dapat digunakan untuk mengatasi pelanggaran asumsi normalitas pada data observasi (Eze et al., 2021). Penelitian ini bertujuan untuk menganalisis pengaruh faktor-faktor sosial ekonomi terhadap persentase kepemilikan rumah di Indonesia dengan pendekatan regresi polinomial berbasis hasil transformasi PCA. Temuan dari penelitian ini diharapkan dapat memberikan masukan berbasis data bagi pemerintah maupun pemangku kepentingan lainnya dalam merumuskan kebijakan perumahan yang lebih tepat sasaran.

2 METODE

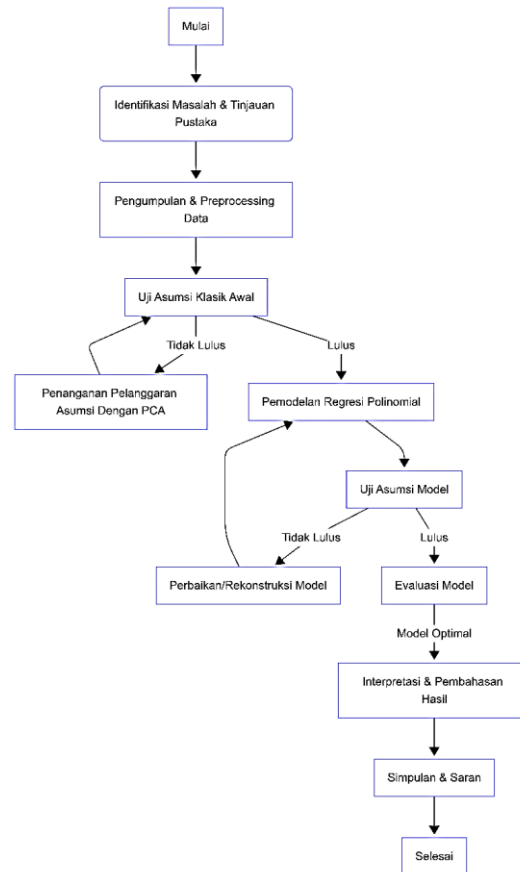
2.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) tahun 2023. Terdapat 1 variabel dependen (Y) dan 5 variabel independen (X_k). Dengan masing-masing variabel dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Nama Variabel	Keterangan Variabel	Skala Pengukuran
Y	Persentase Kepemilikan Rumah	Persentase Rumah Tangga menurut Provinsi, Klasifikasi Desa, dan Status Kepemilikan Bangunan Tempat Tinggal yang Ditempati Milik Sendiri (Persen), 2023	Rasio
X ₁	Kepadatan Penduduk (per km ²)	Kepadatan Penduduk Menurut Provinsi, 2023	Rasio
X ₂	Laju PDRB per kapita (dalam persentase)	Laju Pertumbuhan Produk Domestik Regional Bruto per Kapita Atas Dasar Harga Konstan 2010 Menurut Provinsi (persen), 2023	Rasio
X ₃	Pengangguran (dalam persentase)	Tingkat Pengangguran Terbuka Menurut Provinsi (Persen), Agustus 2023	Rasio
X ₄	Penduduk Miskin (dalam persentase)	Persentase Penduduk Miskin (P0) Menurut Provinsi dan Daerah (Persen), Maret 2023	Rasio
X ₅	<i>Gini Ratio</i>	<i>Gini Ratio</i> Menurut Provinsi dan Daerah, Maret 2023	Rasio

Dalam penelitian ini, seluruh variabel akan menggunakan data agregat tingkat provinsi yang mencakup wilayah perkotaan dan perdesaan. Hal ini dilakukan untuk menjaga konsistensi data, mengingat sebagian besar variabel hanya tersedia dalam bentuk gabungan, sementara hanya beberapa yang memiliki pemisahan antara perkotaan dan perdesaan.



Gambar 1. Skema Metode Penelitian

2.2 Preprocessing Data

Pada tahapan *preprocessing* data, dilakukan pemilihan dan pembersihan data yaitu pembuangan data observasi yang kosong (Rahman et al., 2024). Proses ini dilakukan menggunakan *Microsoft Excel*. Dalam hal ini, terdapat data provinsi yang kosong yaitu Papua Barat Daya, Papua Selatan, Papua Tengah dan Papua Pegunungan. Setelah dilakukan pembersihan pada data kosong, tersisa 34 Provinsi yang akan dilanjutkan ke proses selanjutnya.

2.3 Outlier

Outlier atau pencilan merupakan nilai yang menyimpang jauh dari pengamatan lain atau di luar batas yang dihitung dari nilai kuartil dan interkuartil data. Suatu data dikatakan *outlier* apabila nilai lebih kecil dari $Q1 - 1,5 \times IQR$ dan lebih besar dari $Q3 + 1,5 \times IQR$ (Sihombing et al., 2023). Batas tersebut disebut dengan *whiskers*. Adanya *outlier* terbukti mengakibatkan tidak terpenuhinya asumsi normalitas (Ohyver, 2013). Sehingga, penting dilakukan penanganan terhadap data *outlier*.

2.4 Multikolinearitas

Multikolinearitas terjadi apabila terdapat keterkaitan yang kuat pada sebagian variabel di dalam model (Yunianto, 2021). Kemudian uji multikolinearitas terhadap variabel-variabel sosial ekonomi (X_k) dilakukan untuk menguji apakah terdapat korelasi yang kuat antara variabel independen dan keadaan multikolinearitas bisa dilihat dari nilai VIF (*Variance Inflation Factor*) (Santoso, 2018). Variabel dengan $VIF > 5$, dijelaskan oleh variabel-variabel independen lainnya. Karena adanya multikolinearitas dapat memberikan dampak serius, maka perlu dilakukan langkah perbaikan untuk mengatasi masalah tersebut dan memastikan nilai variabel dependen yang diprediksi lebih akurat. Jika salah satu variabel independen mengalami masalah multikolinearitas, maka variabel tersebut bisa dihapus dari analisis. Namun, jika

banyak variabel yang memiliki VIF lebih besar dari 5, teknik seperti *Principal Component Analysis* (PCA) bisa digunakan untuk mengatasi masalah multikolinearitas tersebut. (Suvarnapathaki, 2023).

$$VIF_k = \frac{1}{1 - R_k^2} \quad (1)$$

Di mana R_k^2 merupakan koefisien determinasi dari variabel independen k, dengan $k = 1, 2, \dots, k$

Selain dengan menggunakan VIF, multikolinearitas untuk banyak variabel dapat dilihat dengan matriks R / matriks korelasi antar variabel (Sriningsih et al., 2018). Koefisien korelasi yang paling umum digunakan dalam matriks korelasi adalah koefisien korelasi Pearson (r), yang mengukur kekuatan dan arah hubungan linear antara dua variabel.

$$r_k = \frac{n(\sum x_k y) - (\sum x_k)(\sum y)}{\sqrt{[n \sum x_k^2 - (\sum x_k)^2][n \sum y^2 - (\sum y)^2]}} \quad (2)$$

Di mana r_k adalah koefisien korelasi Pearson dari variabel independen k, dengan $k = 1, 2, \dots, k$

Tabel 2. Interpretasi *Guilford* Koefisien Korelasi

Nilai Mutlak dari R	Tingkat Hubungan
< 0,19	Hampir Tidak Ada
0,20 - 0,39	Rendah
0,40 - 0,69	Sedang
0,70 - 0,89	Kuat
0,90 - 1,00	Sangat Kuat
$\geq 0,30$	Signifikan

Sumber: (Aswegen & Engelbrecht, 2009)

Pada Tabel 2 ditampilkan interpretasi *Guilford* untuk tingkat hubungan berdasarkan nilai mutlak dari R, apabila nilai R negatif mengindikasikan arah hubungan yang berlawanan. Nilai R dengan nilai sama dengan atau lebih besar dari 0,70 memiliki hubungan yang kuat. Apabila terdapat 5 variabel independen yang saling berhubungan kuat maka terjadi multikolinearitas yang perlu diwaspadai.

2.5 *Principal Component Analysis* (PCA)

Menurut Ramadhan et al. (2024), metode *Principal Component Analysis* (PCA) pertama kali diperkenalkan oleh Harold Hotelling. PCA merupakan teknik reduksi dimensi yang berfungsi untuk menghilangkan multikolinearitas dari variabel independen. PCA dapat diterapkan untuk mengatasi masalah multikolinearitas dalam analisis regresi. Dengan membentuk komponen utama dari variabel independen yang mengandung multikolinearitas, dapat dilihat variabel independen mana yang memiliki pengaruh signifikan dan variabel independen mana yang dapat diidentifikasi sebagai variabel yang penting (Ramadhan et al., 2024). Selain itu, pendekatan PCA dapat digunakan untuk mengatasi pelanggaran asumsi normalitas misalnya terdapat variabel observasi yang tidak berdistribusi normal (Eze et al., 2021). Variansi dari komponen utama sama dengan nilai eigen yang terkait dengannya. Hal ini dikenal sebagai *Kaiser's Criterion*. Metode lain untuk menentukan jumlah komponen utama (*Principal Components* atau PC) yang tepat untuk digunakan adalah dengan melihat *scree plot* dan mengamati adanya siku atau *elbow*. (Suvarnapathaki, 2023). Setelah informasi dari variabel independen dilakukan reduksi dimensi menjadi beberapa komponen (misalnya FAC1 dan

FAC2). Komponen utama diinterpretasi berdasarkan *loading* (kontribusi masing-masing variabel independen asal).

2.6 Regresi Polinomial

Regresi polinomial merupakan jenis regresi khusus yang bekerja pada hubungan lengkung (*curvilinear*) antara variabel dependen dan variabel independen (Eka et al., 2021). Regresi polinomial adalah model regresi yang dipangkatkan meningkat sampai orde ke- n untuk menangkap kecocokan model dengan data. Saat data yang digunakan tidak memenuhi syarat uji linearitas, regresi polinomial cocok dilakukan.

Bentuk umum regresi polinomial adalah sebagai berikut: (Malensang et al., 2013)

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_{nk} X_k^{2n} + \epsilon \quad (3)$$

Keterangan:

Y = variabel dependen

X = variabel independen

$\beta_0, \beta_1, \dots, \beta_{nk}$ = koefisien regresi, dengan $k = 1, 2, \dots, k$ dan $n = 1, 2, \dots, n$

ϵ = *error* (galat)

Jika terdapat lebih dari satu variabel independen, maka persamaan regresi polinomial akan mencakup seluruh variabel tersebut beserta pangkat-pangkatnya sesuai dengan derajat polinomial yang ditentukan.

2.7 Principal Component Regression (PCR)

Principal Component Regression (PCR) merupakan teknik regresi yang menggabungkan pendekatan *Principal Component Analysis* (PCA) dengan metode regresi secara umum (Suvarnapathaki, 2023). PCR digunakan untuk mengatasi permasalahan multikolinearitas dengan terlebih dahulu mentransformasikan variabel-variabel independen yang saling berkorelasi menjadi sejumlah komponen utama yang bersifat tidak berkorelasi, kemudian komponen-komponen tersebut digunakan dalam pemodelan regresi.

2.8 Kriteria Evaluasi Model

Pemilihan model terbaik dilakukan dengan mempertimbangkan nilai koefisien determinasi (*R square*), *adjusted R square*, serta hasil uji signifikansi dari model regresi (Eka et al., 2021). Model dengan kecocokan yang baik terhadap data, namun tetap menjaga kesederhanaan model, akan dijadikan sebagai model akhir yang digunakan dalam interpretasi hasil.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (4)$$

Di mana y_i adalah nilai aktual, $\hat{y}_i$ adalah nilai taksiran dari model regresi, dan $\bar{y}$ adalah nilai rata rata dari y , dengan $i = 1, 2, \dots, n$

$$Adj R^2 = 1 - (1 - R^2) \frac{n - 1}{n - k - 1} \quad (5)$$

Dengan keterangan n adalah jumlah observasi (ukuran sampel) dan k adalah jumlah variabel independen dalam model.

Kedua nilai ini digunakan untuk meninjau seberapa besar variasi variabel dependen dapat dijelaskan oleh variabel independen dalam model, dengan *Adjusted R Square* digunakan sebagai salah satu indikator untuk melihat potensi *overfitting* atau tidaknya suatu model regresi. Batas nilai *R square* antara 0,50 hingga 0,99 dapat diterima dalam penelitian ilmu sosial, terutama ketika sebagian besar variabel independen signifikan secara statistik. Namun, satu hal

yang perlu diperhatikan adalah bahwa nilai *R square* yang tinggi tersebut tidak boleh disebabkan oleh hubungan sebab-akibat yang palsu (*spurious causation*) atau oleh multikolinearitas antar variabel penjelas (Ozili, 2023). Sehingga, apabila tidak terdapat multikolinearitas dan variabel independen signifikan secara statistik, 0,50 menjadi batas minimal *R square* dalam penelitian ilmu sosial.

3 HASIL DAN PEMBAHASAN

3.1 Analisis Statistika Deskriptif

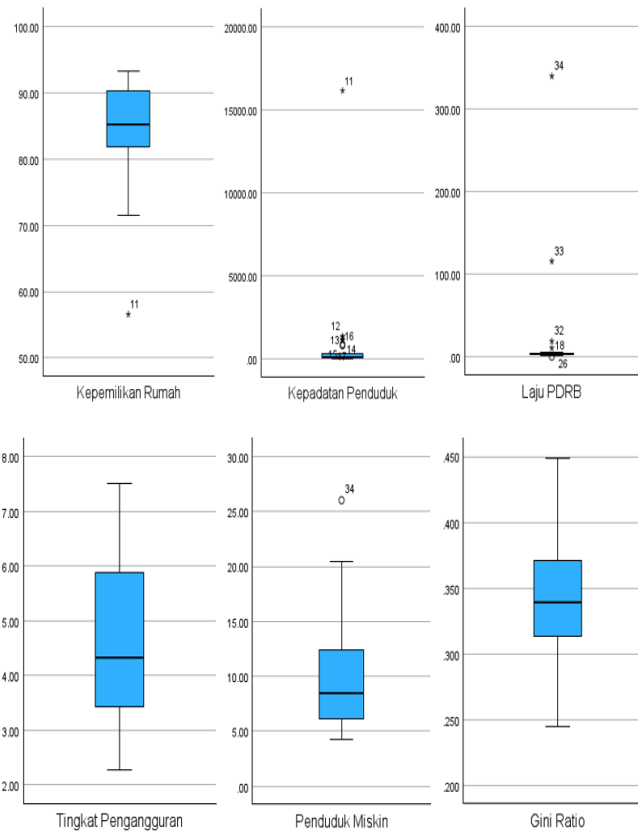
Setelah dilakukan *cleaning* data melalui *software* Microsoft Excel untuk menghapus data provinsi yang kosong dan dilakukan analisis statistika deskriptif, hasilnya disajikan dalam bentuk tabel 3. Median atau Q2 merupakan nilai tengah dari data, kuartil 1 atau Q1 adalah nilai yang membagi 25% data terbawah, dan kuartil 3 atau Q3 adalah nilai yang membagi 75% data teratas. Dalam boxplot, area antara Q1 hingga Q3 digambarkan dengan kotak berwarna biru dan mewakili sebaran utama data yang disebut dengan *interquartile range* (IQR). Sedangkan garis tipis yang menjulur dari kedua ujung kotak merupakan batas nilai normal data atau disebut dengan *Whiskers*. *Whiskers* dihitung berdasarkan $Q1 - 1.5 \times IQR$ dan $Q3 + 1.5 \times IQR$. Jika terdapat titik-titik data yang berada di luar batas tersebut, maka data ini disebut *outlier* atau pencilan. Kemudian hasil analisis statistika deskriptif itu disajikan ke dalam bentuk boxplot agar lebih mudah diinterpretasikan.

Tabel 3. Data Kepemilikan Rumah dan Indikator Sosial Ekonomi di Indonesia Tahun 2023

	Y	X1	X2	X3	X4	X5
Rerata Data	84,04	752,65	17,34	4,61	10,09	0,344
Median	85,275	102,5	3,62	4,32	8,425	0,339
Standar Deviasi	7,58	2708,4	59,32	1,398	5,11	0,046
Q1	82,175	56	3,19	3,4875	6,24	0,314
Q3	89,8575	272,75	4,3475	5,7625	12,2525	0,37075
IQR	7,6825	216,75	1,1575	2,275	6,0125	0,05675

Sumber data: Badan Pusat Statistika (2023)

Kumpulan *boxplot* yang ditampilkan pada Gambar 2 menggambarkan masing-masing sebaran data dari variabel independen dan dependen. Pada variabel dependen kepemilikan rumah dan variabel independen penduduk miskin terdapat 1 *outlier*, variabel independen kepadatan penduduk dan laju PDRB terdapat cukup banyak *outlier*, serta pada variabel independen tingkat pengangguran dan gini ratio tidak terdapat *outlier* sama sekali.



Gambar 2. *Boxplot* Variabel Dependen dan Independen

Tabel 4. Uji Normalitas dengan *Shapiro-Wilk*

	Shapiro-Wilk			Berdistribusi Normal
	<i>Statistic</i>	<i>df</i>	<i>Sig. (p-value)</i>	
Kepemilikan Rumah	0,863	34	<0,001	Tidak
Kepadatan Penduduk	0,255	34	<0,001	Tidak
Laju PDRB	0,260	34	<0,001	Tidak
Tingkat Pengangguran	0,964	34	0,321	Ya
Penduduk Miskin	0,888	34	0,002	Tidak
Gini Ratio	0,981	34	0,803	Ya

Tabel 4 menunjukkan hasil uji Shapiro-Wilk yang dilakukan dalam pengujian normalitas. Shapiro-Wilk dipilih karena sampel yang digunakan di bawah 50, tepatnya ada 34 observasi. Kriteria keputusan pada uji ini adalah apabila nilai signifikansi (Sig.) lebih besar dari 0,05 yang berarti data berdistribusi normal dan data tidak berdistribusi normal apabila nilai signifikansi (Sig.) lebih kecil dari 0,05. Antara boxplot dan hasil uji Shapiro-Wilk berkaitan dan memberikan informasi bahwa data tanpa *outlier* cenderung berdistribusi normal dan data dengan *outlier* (walaupun hanya terdapat satu *outlier*) berdistribusi tidak normal. Hal ini menunjukkan uji Shapiro-Wilk rentan terhadap *outlier*.

Nilai-nilai dan variabel yang tidak berdistribusi normal tidak dihapus karena mencerminkan kondisi riil yang relevan dan penting untuk dianalisis. Oleh karena itu, *outlier* pada data observasi dan variabel yang tidak berdistribusi normal tetap dipertahankan dalam model untuk menjaga keutuhan informasi yang dibawa oleh data tersebut. Selanjutnya dianalisis matriks korelasi, yang dirangkum pada Tabel 5. Tabel ini menunjukkan koefisien korelasi Pearson (r_k) antarvariabel dan sekaligus menjadi indikator awal untuk mendeteksi

potensi masalah multikolinearitas di antara variabel-variabel independen sebelum pemodelan regresi.

Tabel 5. Matriks Korelasi

	X1	X2	X3	X4	X5
X1	1				
X2	-0.060	1			
X3	0.284	-0.203	1		
X4	-0.217	0.626	-0.35	1	
X5	0.404	0.175	-0.008	0.191	1

Tabel 6. Matriks Korelasi Variabel Kuadrat (Orde-2)

	X1	X2	X3	X4	X5	X1 ²	X2 ²	X3 ²	X4 ²	X5 ²
X1	1									
X2	-0.060	1								
X3	0.284	-0.203	1							
X4	-0.217	0.626	-0.35	1						
X5	0.404	0.175	-0.008	0.191	1					
X1 ²	0.991	-0.038	0.243	-0.194	0.333	1				
X2 ²	-0.053	0.977	-0.231	0.582	0.167	-0.035	1			
X3 ²	0.300	-0.176	0.988	-0.319	0.050	0.252	-0.196	1		
X4 ²	-0.166	0.773	-0.329	0.969	0.172	-0.138	0.732	-0.296	1	
X5 ²	0.428	0.167	0.005	0.176	0.996	0.355	0.161	0.062	0.159	1

Berdasarkan matriks korelasi antar variabel independen asli, tidak ditemukan adanya multikolinearitas yang perlu diwaspadai. Namun, dalam pemodelan regresi polinomial, digunakan variabel turunan seperti X_1^2 , yang secara matematis memiliki hubungan kuat dengan variabel aslinya. Transformasi ini menimbulkan multikolinearitas baru antara variabel asli dan hasil kuadratnya, yang terdeteksi melalui nilai koefisien determinasi (R^2) mendekati 1 pada beberapa variabel-variabel tersebut, serta terdapat hubungan yang cenderung kuat antara X_2 dan X_4^2 . Oleh karena itu, digunakan *Principal Component Analysis* (PCA) untuk mengurangi dimensi dan menghilangkan multikolinearitas, dengan cara mentransformasikan variabel-variabel polinomial menjadi kombinasi linier baru yang saling orthogonal (tidak berkorelasi). Dengan pendekatan ini, model regresi tetap dapat mempertahankan informasi variasi dari variabel aslinya tanpa terpengaruh oleh multikolinearitas, sehingga hasil estimasi menjadi lebih stabil dan interpretatif.

3.2 *Principal Component Analysis* (PCA)

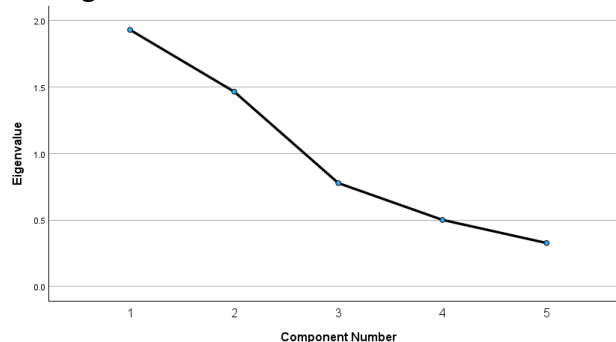
Pada penelitian ini, penulis menggunakan *software* IBM SPSS untuk melakukan analisis data. Dari hasil analisis PCA didapatkan *Total Variance Explained* pada Tabel 7 berikut. *Total Variance Explained* menampilkan jumlah komponen dan *initial eigenvalues* baik total dan kumulatif. Ketiga nilai ini memberikan informasi mengenai jumlah terbaik komponen yang akan mewakili kelima variabel independen yang direduksi.

Tabel 7. *Total Variance Explained*

Component	Initial Eigenvalues
-----------	---------------------

	<i>Total</i>	<i>Cumulative %</i>
1	1,930	38,606
2	1,465	67,910
3	0,777	83,459
4	0,501	93,470
5	0,326	100,000

Berdasarkan *Total Variance Explained* pada Tabel 3, nilai *eigenvalue* komponen ke-1 dan ke-2 lebih besar dari 1, sedangkan komponen ke-3 dan selanjutnya lebih kecil dari 1. Hal ini sesuai dengan *Kaiser's Criterion*. Sehingga, jumlah komponen faktor yang diambil sebanyak 2 dan 2 komponen faktor tersebut mampu menjelaskan 67,91% variasi data berdasarkan nilai *Cumulative*. Gambar 3 menampilkan *scree plot* yang memetakan *eigenvalue* masing-masing komponen pada sumbu vertikal terhadap urutan komponen pada sumbu horizontal. Visualisasi ini berfungsi sebagai alat diagnostik utama untuk menentukan jumlah komponen yang layak dipertahankan, dengan cara mengidentifikasi titik “siku” (elbow) yaitu transisi tajam dari penurunan curam ke kemiringan linier.



Gambar 3. *Scree Plot*

Pada Gambar 3, terlihat adanya siku atau *elbow* pada *component number* ke-2 dan menandakan *scree plot* mendukung interpretasi dari tabel *total variance explained*. Artinya setelah komponen ke-2, tambahan komponen tidak banyak menambah informasi dan 2 komponen faktor sudah cukup mewakili. Tabel *Component Matrix* menampilkan masing-masing variabel independen yang direduksi serta nilai *loading* yang diwakili oleh komponen. Nilai *loading* yang dominan menandakan informasi variabel independen terdapat dimensi komponen tersebut. Negatif dalam nilai *loading* dapat diinterpretasikan bahwa nilai yang diwakilkan memberikan informasi hubungan berlawanan arah.

Tabel 8. *Component Matrix*

	<i>Component</i>	
	1	2
Kepadatan Penduduk	-0,369	0,791
Laju PDRB	0,782	0,259
Pengangguran	-0,615	0,264
Penduduk Miskin	0,879	0,122
Gini Ratio	0,173	0,830

Interpretasi komponen PCA dari tabel *Component Matrix* pada tabel 8 didasarkan pada nilai terbesar dari masing-masing variabel independen pada komponen faktor yang mewakili. Komponen Faktor 1 (FAC1) didominasi kontribusi dari Laju PDRB (+), Persentase Pengangguran (-), dan Kemiskinan (+). Komponen ini ditafsirkan sebagai dimensi kesejahteraan ekonomi. Nilai negatif pada *loading* menyatakan arah berlawanan. Sedangkan

komponen Faktor 2 (FAC2) didominasi kontribusi dari Kepadatan Penduduk (+) dan Gini Ratio (+). Komponen ini ditafsirkan sebagai dimensi demografi dan distribusi pendapatan.

3.3 Analisis Regresi Polinomial

Selanjutnya, dilakukan analisis regresi polinomial Persentase Kepemilikan Rumah menggunakan variabel independen hasil PCA sekaligus pengujian kelayakan model. Dilakukan beberapa kali iterasi dalam analisis regresi polinomial untuk menghasilkan model terbaik berdasarkan uji kelayakan model.

3.4 Uji Kelayakan Model

i. Model Awal Regresi Polinomial

Untuk menunjukkan hubungan antara masing-masing komponen faktor terhadap variabel dependen Kepemilikan Rumah, digunakan analisis regresi polinomial yang membentuk model estimasi. Model estimasi awalnya adalah:

$$Y = 86,204 + 5,478 \text{ FAC1} - 1,396 \text{ FAC1}^2 + 0,351 \text{ FAC2} - 0,829 \text{ FAC2}^2 \quad (6)$$

Uji kelayakan model dilakukan menggunakan *software* IBM SPSS dengan masing-masing penjelasan pada poin selanjutnya.

1. Uji Serentak (Uji F)

Uji serentak dilakukan dengan uji F (uji *Analysis of Variance*/ANOVA), yang bertujuan untuk mengetahui apakah secara bersama-sama variabel independen berpengaruh signifikan terhadap variabel dependen. Dengan H_0 menyatakan bahwa semua koefisien regresi variabel independen sama dengan nol pada interval konfidensi 95%, artinya tidak ada pengaruh secara serentak terhadap variabel dependen. Sebaliknya, hipotesis alternatif menyatakan bahwa paling tidak terdapat satu variabel independen yang berpengaruh secara signifikan. Kriteria penolakan H_0 adalah saat F_{hitung} lebih besar dari $F_{(0,05;df \text{ regression},df \text{ residual})}$ atau nilai signifikansi (Sig.) yang lebih kecil 0,05.

Tabel 9. Statistika Uji F: ANOVA

<i>Model</i>		<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig. (p-value)</i>
1	<i>Regression</i>	1096,874	4	274,218	9,281	<0,001 ^b
	<i>Residual</i>	856,830	29	29,546		
	<i>Total</i>	1953,704	33			

Berdasarkan tabel 9, nilai Sig. kurang dari 0,001, $F_{\text{hitung}} = 9,281$ dan $F_{(0,05;4,29)} = 4,04$. Diartikan bahwa nilai signifikansi (Sig.) berada di bawah ambang batas 0,05 dan $F_{\text{hitung}} > F_{\text{tabel}}$, sehingga model secara keseluruhan dianggap signifikan secara statistik pada tingkat kepercayaan 95%. Secara bersama-sama variabel independen FAC1, FAC1², FAC2, FAC2² memiliki pengaruh kepada variabel dependen Kepemilikan Rumah.

2. Uji Parsial (Uji t)

Uji parsial dilakukan dengan uji t, yang bertujuan untuk mengetahui apakah masing-masing variabel independen secara individu berpengaruh signifikan terhadap variabel dependen. Dalam uji ini, hipotesis nol (H_0) menyatakan bahwa koefisien regresi dari satu variabel independen tertentu sama dengan nol pada interval konfidensi 95%, artinya variabel tersebut tidak berpengaruh secara signifikan terhadap variabel dependen. Sebaliknya, hipotesis alternatif menyatakan bahwa variabel tersebut memiliki pengaruh yang signifikan. Kriteria penolakan H_0 adalah jika nilai signifikansi (Sig.) lebih kecil dari 0,05. Jika H_0 ditolak, maka dapat

disimpulkan bahwa variabel independen tersebut berpengaruh secara signifikan terhadap variabel dependen.

Tabel 10. Statistik Uji t

	<i>Model</i>	<i>Unstandardized B</i>	<i>t_{hitung}</i>	<i>Sig. (p-value)</i>	<i>Keputusan</i>
1	(Constant)	86,204	77,28	<0,001	
	FAC1	5,478	3,232	0,003	Signifikan
	FAC2	0,351	0,230	0,820	Tidak Signifikan
	FAC1 ²	-1,396	-2,490	0,019	Signifikan
	FAC2 ²	-0,829	-1,428	0,164	Tidak Signifikan

Nilai signifikansi (Sig.) FAC1 dan FAC1² pada tabel lebih rendah dari 0,05. Untuk nilai signifikansi (Sig.) FAC2 dan FAC2² lebih besar dari 0,05. Oleh karena itu, komponen faktor FAC1 dan bentuk kuadratnya signifikan secara statistik, sedangkan FAC2 dan bentuk kuadratnya tidak signifikan. Maka, FAC2 dipertimbangkan untuk dihapus dari model regresi polinomial dan dilakukan kembali uji secara lengkap pada model regresi baru (model final).

ii. Model Final Regresi Polinomial

Sebagai hasil dari proses iterasi dan perbaikan dari model sebelumnya, analisis regresi final dilakukan dengan hanya menggunakan komponen FAC1 dan kuadratnya sebagai variabel independen untuk menunjukkan hubungan terhadap variabel dependen Kepemilikan Rumah. Model estimasi final yang dihasilkan adalah:

$$Y = 85,971 + 7,4 \text{ FAC1} - 1,985 \text{ FAC1}^2 \quad (7)$$

Selanjutnya uji kelayakan model dilakukan kembali terhadap model final ini.

1. Uji Serentak (Uji F)

Tabel 11. Statistika Uji F: ANOVA

<i>Model</i>		<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig. (p-value)</i>
1	<i>Regression</i>	1007,156	2	503,578	16,492	<0,001 ^b
	<i>Residual</i>	946,548	31	30,534		
	<i>Total</i>	1953,704	33			

Dalam Tabel 11, hasil dari uji serentak yang dilakukan menghasilkan nilai Sig. kurang dari 0,001 dan $F_{hitung} = 16,492 > F_{(0,05;2,31)} = 3,30$. Hal ini berarti nilai signifikansi dan statistik uji F menolak H_0 , sehingga model secara keseluruhan dianggap signifikan secara statistik pada interval konfidensi 95%. Variabel independen FAC1 dan FAC1² secara serentak memiliki pengaruh kepada variabel dependen Kepemilikan Rumah tahun 2023 di Indonesia.

2. Uji Parsial (Uji t)

Tabel 12. Statistik Uji t

	<i>Model</i>	<i>Unstandardized B</i>	<i>t</i>	<i>Sig. (p-value)</i>	<i>Keputusan</i>
1	(Constant)	85,971	82,64	<0,001	
	FAC1	7,4	5,673	<0,001	Signifikan
	FAC1 ²	-1,985	-4,491	<0,001	Signifikan

Pada hasil uji parsial di Tabel 12, nilai signifikansi (Sig.) FAC1 dan FAC1² lebih rendah dari 0,05. Oleh karena itu, komponen faktor FAC1 dan bentuk kuadratnya signifikan secara statistik. Sehingga, uji dilanjutkan ke uji kelayakan model lainnya.

3. Koefisien Determinasi (R^2)

Tabel *Model Summary* pada Tabel 13 menyajikan nilai koefisien determinasi (*R Square*) dan *Adjusted R Square*.

Tabel 13. Model Summary

<i>Model</i>	<i>R Square</i>	<i>Adjusted R Square</i>
1	0,516	0,484

Pada Tabel 13 menyatakan koefisien determinasi (R^2) dari regresi polinomial komponen faktor PCA adalah sebesar 0,516 yang artinya 51,6% variasi pada variabel Kepemilikan Rumah dapat dijelaskan oleh keempat independen yaitu FAC1 dan FAC1². Sedangkan Adjusted R^2 sebesar 0,484 menunjukkan model masih tergolong netral atau tidak terlalu overfit.

4. Asumsi klasik IIDN

Untuk memastikan kualitas estimasi parameter model, penting bagi model regresi untuk memiliki sifat *BLUE* (*Best Linear Unbiased Estimator*). Selain itu, asumsi bahwa residual model regresi bersifat IIDN (identik, independen, dan berdistribusi normal) juga harus terpenuhi (Fitriyah et al., 2021)

a. Uji Normalitas

Hasil uji normalitas menggunakan uji Shapiro-Wilk disajikan pada Tabel 13. Kriteria keputusan pada uji ini adalah apabila nilai signifikansi (Sig.) lebih besar dari 0,05, maka data diasumsikan berdistribusi normal, dan sebaliknya apabila lebih kecil dari 0,05, data tidak berdistribusi normal.

Tabel 14. Uji Normalitas dengan Shapiro-Wilk

	Shapiro-Wilk		
	Statistic	df	Sig. (<i>p-value</i>)
<i>Unstandardized Residual</i>	0,975	34	0,601

Berdasarkan Tabel 14, hasil uji Shapiro-Wilk untuk *Unstandardized Residual* menunjukkan nilai signifikansi (Sig.) sebesar 0,601. Karena nilai signifikansi yang diperoleh (0,601) lebih besar dari ambang batas 0,05, maka dapat disimpulkan bahwa residual model berdistribusi normal. Kondisi ini memenuhi asumsi normalitas dalam analisis regresi, sehingga inferensi statistik dari model dapat dianggap valid.

b. Uji Multikolinearitas

Uji multikolinearitas dilakukan untuk mengecek ada tidaknya hubungan linier yang tinggi antarvariabel independen dalam model regresi, yang dapat menyebabkan estimasi koefisien yang tidak stabil. Uji ini dilakukan dengan menganalisis nilai *Tolerance* dan *Variance Inflation Factor* (*VIF*). Hasil uji multikolinearitas disajikan pada Tabel 15.

Tabel 15. Coefficients

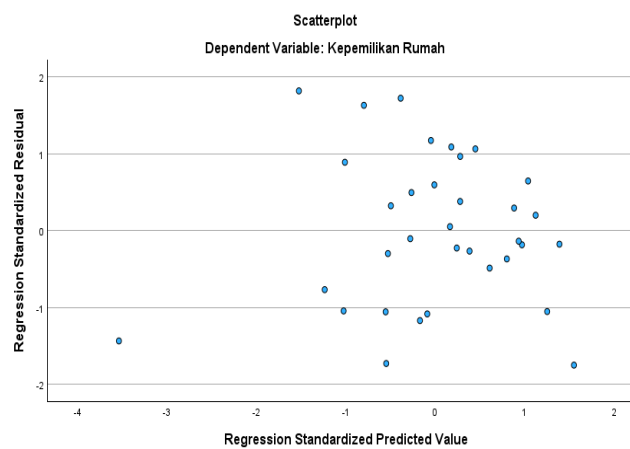
<i>Model</i>	<i>Collinearity Statistics</i>	
	<i>Tolerance</i>	<i>VIF</i>
1 (Constant)		
FAC1	0,544	1,839

$$\frac{\text{FAC1}^2}{0,544} \quad 1,839$$

Berdasarkan Tabel 15, nilai *Tolerance* untuk variabel FAC1 dan FAC1^2 adalah 0,544, sementara nilai VIF untuk kedua variabel tersebut adalah 1,839. Karena nilai *VIF* berada di bawah batas 5 (atau sering juga menggunakan batas 10) dan nilai *Tolerance* di atas 0,1, hal ini menunjukkan bahwa tidak terjadi kondisi multikolinearitas yang perlu diwaspadai antarvariabel independen dalam model. Hal ini mendukung persyaratan penerimaan nilai 0,5 sebagai batas minimal pada *R square* pada penelitian ini. Sehingga, secara tidak langsung model regresi polinomial dengan komponen faktor yang diuji cocok digunakan untuk menjelaskan pengaruh variabel-variabel sosial ekonomi terhadap kepemilikan rumah.

c. Uji Homoskedastisitas

Uji homoskedastisitas dilakukan dengan bantuan *scatterplot* antara nilai *Regression Standardized Residual* dan *Regression Standardized Predicted Value*. Dikatakan terjadinya heteroskedastisitas apabila pola sebaran membentuk suatu pola sistematis dan tidak acak.



Gambar 4. Uji Homoskedastisitas dengan *Scatterplot* (Visual)

Pola sebaran titik pada *scatterplot* terlihat acak dan tidak membentuk pola tertentu seperti kipas, garis melengkung, atau bentuk sistematis lainnya. Titik-titik menyebar secara merata di atas dan bawah sumbu nol, tanpa pola melebar atau menyempit seiring meningkatnya nilai prediksi. *Scatterplot* tersebut tidak menunjukkan indikasi adanya heteroskedastisitas. Artinya, ragam residual relatif konstan di seluruh nilai prediksi, sehingga asumsi homoskedastisitas (varian error yang konstan) dipenuhi.

Pada hasil analisis regresi, didapatkan persamaan regresi terbaik dalam penelitian ini yang ditulis pada persamaan (8):

$$Y = 85,971 + 7,4 \text{ FAC1} - 1,985 \text{ FAC1}^2 \quad (8)$$

Berdasarkan hasil regresi polinomial orde 2 yang menggunakan faktor pertama (FAC1) hasil analisis faktor utama (*Principal Component Analysis*) beserta bentuk kuadratnya (FAC1^2) terhadap variabel dependen Kepemilikan Rumah, diperoleh model dengan nilai *R square* sebesar 0,516 dan *Adjusted R square* sebesar 0,484. Artinya, sekitar 51,6% variabilitas dalam kepemilikan rumah antar provinsi di Indonesia pada tahun 2023 dapat dijelaskan oleh komponen faktor pertama (FAC1) dan bentuk kuadratnya (FAC1^2). Selanjutnya, interpretasi koefisien regresi menunjukkan bahwa faktor pertama (FAC1) yang merepresentasikan variabel-variabel seperti laju PDRB per kapita, persentase pengangguran, dan persentase kemiskinan, memiliki pengaruh positif dan signifikan terhadap kepemilikan rumah ($\beta = 7,4$). Hal ini menunjukkan bahwa peningkatan pada kondisi yang direpresentasikan oleh faktor ini

(misalnya tingginya laju PDRB, rendahnya tingkat pengangguran, dan tingginya tingkat kemiskinan) cenderung meningkatkan persentase kepemilikan rumah. Namun demikian, $FAC1^2$ juga menunjukkan pengaruh negatif dan signifikan ($\beta = -1,985$), mengindikasikan adanya hubungan nonlinier berbentuk parabola terbuka ke bawah. Interpretasi ini dapat dimaknai bahwa peningkatan $FAC1$ memang awalnya meningkatkan kepemilikan rumah, tetapi setelah melewati titik tertentu, efeknya cenderung menurun. Ini mencerminkan adanya titik optimal kesejahteraan ekonomi di mana kepemilikan rumah mencapai puncaknya, kemudian cenderung stagnan atau bahkan turun.

4 KESIMPULAN

Penelitian ini menunjukkan bahwa indikator sosial ekonomi secara bersama-sama berpengaruh terhadap tingkat kepemilikan rumah di Indonesia, dengan pola hubungan yang tidak linier. Melalui pendekatan *Principal Component Analysis* (PCA), lima indikator sosial ekonomi yang saling berkorelasi direduksi menjadi dua komponen utama. Regresi polinomial orde dua menunjukkan bahwa komponen pertama ($FAC1$), yang merepresentasikan kombinasi dari PDRB, pengangguran, dan kemiskinan, memiliki pengaruh positif dan signifikan terhadap kepemilikan rumah. Komponen kedua ($FAC2$), yang berkaitan dengan kepadatan penduduk dan gini ratio, tidak berpengaruh signifikan. Sehingga komponen kedua dihilangkan dari model. Model final dengan $FAC1$ menjelaskan lebih dari separuh variabilitas kepemilikan rumah ($R\text{ square} = 0,516$), dan mengindikasikan pentingnya mempertimbangkan pola tersembunyi variabel sosial ekonomi serta hubungan nonlinier dalam pemodelan. Di sisi lain, kemiskinan menunjukkan hubungan positif terhadap kepemilikan rumah. Temuan ini membuka peluang untuk kajian lebih lanjut guna memahami dinamika sosial ekonomi yang mendasari hubungan tersebut. Hal ini memiliki relevansi penting dalam perumusan kebijakan perumahan bahwa upaya peningkatan kepemilikan rumah tidak cukup hanya dengan memperbaiki satu aspek ekonomi saja, melainkan perlu pendekatan yang lebih menyeluruh dan seimbang terhadap berbagai faktor sosial ekonomi yang memengaruhi akses terhadap kepemilikan rumah. Penelitian selanjutnya disarankan untuk memperluas cakupan variabel, menambahkan dimensi harga rumah dan status migrasi penduduk, serta menggunakan data panel untuk menggambarkan dinamika jangka panjang ataupun menggunakan metode analisis lain yang lebih menangkap informasi tersebut.

UCAPAN TERIMA KASIH

Apresiasi kami sampaikan kepada Badan Pusat Statistik yang telah menyediakan data terbuka mengenai variabel sosial ekonomi di Indonesia. Data-data tersebut sangat berharga dan menjadi fondasi penting dalam penelitian ini. Terima kasih juga kami haturkan khusus kepada dosen Program Studi Statistika Universitas Terbuka yaitu Ibu Ika Nur Laily Fitriana yang telah memberikan banyak dukungan, bimbingan, masukan yang membangun dan ilmu pengetahuan yang sangat bermanfaat sehingga penulis dapat menyelesaikan artikel ini.

DAFTAR PUSTAKA

- Aswegan, A. S. V., & Engelbrecht, A. S. (2009). The relationship between transformational leadership, integrity and an ethical climate in organisations. *SA Journal of Human Resource Management*, 7(1). <https://doi.org/10.4102/sajhrm.v7i1.175>
- Bank Indonesia. (2023). *Residential Property Price Survey Q4/2022: Residential Property Prices Continue to Rise*. https://www.bi.go.id/en/publikasi/ruang-media/news-release/Pages/sp_253923.aspx
- CIMB Niaga. (2021). *Faktor Pembeda Harga Rumah dan Rata-rata Harga Rumah*. <https://www.cimbniaga.co.id/id/inspirasi/perencanaan/faktor-pembeda-harga-rumah-dan-rata-rata-harga-rumah>

- Eka, A., Juarna, A., Informatika, T., Industri, F. T., & Gunadarma, U. (2021). Prediksi Produksi Daging Sapi Nasional dengan Metode Regresi Linier dan Regresi Polinomial. *Jurnal Ilmiah Komputasi*, 20(2), 209–215. <https://doi.org/10.32409/jikstik.20.2.2722>
- Eze, N. M., Asogwa, O. C., & Eze, C. M. (2021). Principal Component Factor Analysis of Some Development Factors in Southern Nigeria and Its Extension to Regression Analysis. *Journal of Advances in Mathematics and Computer Science*, 36(3), 132–160. <https://doi.org/10.9734/jamcs/2021/v36i330351>
- Fitriyah, Z., Irsalina, S., K, A. R. H., & Widodo, E. (2021). Analisis Faktor Yang Berpengaruh Terhadap Ipm Menggunakan Regresi Linear Berganda. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 2(3), 282–291. <https://doi.org/10.46306/lb.v2i3.86>
- Hafidzi, A. H., Kamilia, I., & Afroh, F. (2024). Homeownership in Low-Income Communities in Indonesia : The Role of Poverty Levels , Inflation , and Bank Indonesia Rate. *JKMP(JurnalKebijakan dan Manajemen Publik)*, 12(2), 122-134. <https://doi.org/10.21070/jkmp.v12i2.1771>
- Kim, J.-S., & Kim, J.-M. (2017). The Relation between Housing Needs and Housing Function according to the Maslow's Theory of Needs. *KIEAE Journal*, 17(4), 13–19. <https://doi.org/10.12813/kieae.2017.17.4.013>
- Malensang, J. S., Komalig, H., & Hatidja, D. (2013). Pengembangan Model Regresi Polinomial Berganda Pada Kasus Data Pemasaran. *Jurnal Ilmiah Sains*, 12(2), 149. <https://doi.org/10.35799/jis.12.2.2012.740>
- Mire, M. S. (2024). Sociodemographic Characteristics Mediated by Poverty as Determinants of Homeownership in Indonesia. *International Journal of Research and Scientific Innovation*, 11(4), 1018–1033. <https://doi.org/10.51244/IJRSI.2024.1104073>
- Natalia, V. V., & Wunas, S. (2016). Housing Development: Sign and Symptoms in Middle City of Indonesia. *Procedia - Social and Behavioral Sciences*, 227, 278–285. <https://doi.org/10.1016/j.sbspro.2016.06.072>
- Ohyver, M. (2013). Penerapan Metode Transformasi Logaritma Natural dan Partial Least Squares Untuk Memperoleh Model Bebas Multikolinier dan Outlier. *Jurnal Mat Stat*, 13(1), 42–51.
- Ozili, P. K. (2023). The acceptable R-square in empirical modelling for social science research. *Social Research Methodology and Publishing Results: A Guide to Non-Native English Speakers*, 115769, 134–143. <https://doi.org/10.4018/978-1-6684-6859-3.ch009>
- Rahman, R., Sudiarjo, A., & Sumaryana, Y. (2024). Prediksi Upah Minimum Provinsi 10 Tahun Kedepan Dengan Menggunakan Model Polynomial Regression. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(4), 7666–7673. <https://doi.org/10.36040/jati.v8i4.9889>
- Ramadhan, R., Fimba, A. B., Fernandes, A. A. R., Solimun, S., Junianto, F. H., Amanda, D. V., & Sumara, R. (2024). Explore The Determinants of Customers Time to Pay House Ownership Loan on Data with High Multicollinearity with Pca-Cox Regression. *Media Statistika*, 17(2), 117–127. <https://doi.org/10.14710/medstat.17.2.117-127>
- Santoso, A. B. (2018). *Tutorial dan Solusi Pengolahan Data Regresi*. Garuda Mas Sejahtera.
- Satoto, E. B. (2023). Boosting Homeownership Affordability for Low-Income Communities

- in Indonesia. *International Journal of Sustainable Development and Planning*, 18(5), 1365–1376. <https://doi.org/10.21070/jkmhlm.v12i2.17>
- Sihombing, P. R., Suryadiningrat, S., Sunarjo, D. A., & Yuda, Y. P. A. C. (2023). Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya. *Jurnal Ekonomi Dan Statistik Indonesia*, 2(3), 307–316. <https://doi.org/10.11594/jesi.02.03.07>
- Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan Multikolinearitas Dengan Menggunakan Analisis Regresi Komponen Utama Pada Kasus Impor Beras Di Provinsi Sulut. *Jurnal Ilmiah Sains*, 18(1), 18-24. <https://doi.org/10.35799/jis.18.1.2018.19396>
- Suvarnapathaki, S. (2023). Use of Principal Component Analysis in Regression Problem. *International Journal of Research and Analytical Reviews (IJRAR)*, 10(3), 125-135. <https://doi.org/10.1729/Journal.36215>
- Yunianto, D. (2021). Analisis pertumbuhan dan kepadatan penduduk terhadap pertumbuhan ekonomi. *Forum Ekonomi*, 23(4), 688–699. <https://doi.org/10.30872/jfor.v23i4.10233>
- Yustie, R., Kiak, N., & Utami, B. (2024). Faktor-Faktor Ekonomi yang Mempengaruhi Status Kepemilikan Rumah di Indonesia Tahun 2017-2022. *Jurnal Akuntansi: Transparansi dan Akuntabilitas*, 12(2), 99–106.