PERBANDINGAN *PRINCIPAL COMPONENT REGRESSION* DAN REGRESI *RIDGE* PADA ANALISIS FAKTOR-FAKTOR INDEKS PEMBANGUNAN MANUSIA

SKRIPSI

OLEH
RIZKA MAULIDA
NIM. 18610020

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022

PERBANDINGAN *PRINCIPAL COMPONENT REGRESSION* DAN REGRESI *RIDGE* PADA ANALISIS FAKTOR-FAKTOR INDEKS PEMBANGUNAN MANUSIA

SKRIPSI

Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)

Oleh
Rizka Maulida
NIM. 18610020

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022

PERBANDINGAN *PRINCIPAL COMPONENT REGRESSION* DAN REGRESI *RIDGE* PADA ANALISIS FAKTOR-FAKTOR INDEKS PEMBANGUNAN MANUSIA

SKRIPSI

Oleh
Rizka Maulida
NIM. 18610020

Telah Diperiksa dan Disetujui Untuk Diuji
Malang, 21 Juni 2022

Dosen Pembimbing I

Fachrur Rozi, M.Si
NIP. 19800527 200801 1 012

Dosen Pembimbing II

Mohammad Nafie Jauhari, M.Si
NIDT. 19870218 20160801 1 056

Mengetahui,
Ketua Program Studi Matematika

Dr. Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005

PERBANDINGAN *PRINCIPAL COMPONENT REGRESSION* DAN REGRESI *RIDGE* PADA ANALISIS FAKTOR-FAKTOR INDEKS PEMBANGUNAN MANUSIA

SKRIPSI

Oleh
Rizka Maulida
NIM. 18610020

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Tanggal 27 Juni 2022

Ketua Penguji	:	Dr. Sri Harini, M.Si	....
Anggota Penguji I	:	Ria Dhea Layla Nur Karisma, M.Si	....
Anggota Penguji II	:	Fachrur Rozi, M.Si	....
Anggota Penguji III	:	Mohammad Nafie Jauhari, M.Si	....

Mengetahui,
Ketua Program Studi Matematika

Dr. Elly Susanti, S.Pd., M.Sc
NIP. 19741129 200012 2 005

PERNYATAAN KEASLIAN TULISAN

Saya yang bertandatangan di bawah ini:

Nama	: Rizka Maulida
NIM	: 18610020
Program Studi	: Matematika
Fakultas	: Sains dan Teknologi
Judul Skripsi	: Perbandingan *Principal Component Regression* dan Regresi *Ridge* pada Analisis Faktor-Faktor Indeks Pembangunan Manusia

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 21 Juni 2022

Yang membuat pernyataan,

Rizka Maulida

NIM. 18610020

MOTO

“Buatlah tujuan untuk hidup, kemudian gunakan segenap kekuatan untuk mencapainya, kamu pasti berhasil”

Utsman Bin Affan

PERSEMBAHAN

Skripsi ini penulis persembahkan untuk:
Ayahanda Muhammad Uman, Ibunda Juwariyah, dan adik Faizatul Laila Zuhriyah yang senantiasa mendoakan, memberikan semangat, kasih sayang, nasehat, dan selalu ada saat penulis butuhkan sehingga penulis dapat bertahan dan berjuang sampai saat ini.

.

KATA PENGANTAR

Segala puji bagi Allah SWT yang telah melimpahkan rahmat, taufik dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Perbandingan *Principal Component Regression* dan Regresi *Ridge* pada Analisis Faktor-Faktor Indeks Pembangunan Manusia". Shalawat serta salam selalu terlimpahkan kepada Nabi Muhammad saw yang telah menuntun manusia ke jalan yang benar.

Dalam penyusunan skripsi ini tidak lepas dari bimbingan dan arahan dari berbagai pihak baik. Untuk itu ucapan terima kasih dan penghargaan yang sebesar-besarnya penulis sampaikan kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Elly Susanti, S.Pd., M.Sc, selaku ketua Program Studi Matematika, Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Fachrur Rozi, M.Si, selaku dosen pembimbing I dan dosen wali yang telah memberikan bimbingan, arahan, dan saran kepada penulis dengan sabar.
5. Mohammad Nafie Jauhari, M.Si, selaku dosen pembimbing II yang telah memberikan bimbingan dan saran kepada penulis.
6. Dr. Sri Harini, M.Si, selaku ketua penguji yang telah memberikan saran dan ilmunya dalam penulisan skripsi ini.
7. Ria Dhea Layla Nur Karisma, M.Si, selaku anggota penguji I yang telah memberikan saran dan ilmunya dalam penulisan skripsi.
8. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
9. Orang tua dan adik yang telah memberikan dukungan, doa dan semangat kepada penulis sehingga penulis dapat menyelesaikan skripsi ini.
10. Seluruh teman-teman dalam Program Studi Matematika angkatan 2018.

11. Semua pihak yang tidak bisa penulis sebutkan satu per satu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Selain itu, penulis juga berharap semoga skripsi ini dapat bermanfaat bagi penulis dan bagi pembaca. *Aamiin*

Malang, 17 Juni 2022

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR SIMBOL

X	: Variabel bebas
Y	: Variabel terikat
Y_i	: Variabel terikat dengan $i = 1,2,\dots,n$
X_{ij}	: Variabel bebas dengan $i = 1,2,\dots,n$; $j = 1,2,\dots,k$
β_j	: Koefisien regresi ke- j dengan $j = 1,2,\dots,k$
ε_i	: *Error* ke-i dengan $i = 1,2,\dots,n$
$\mathbf{\Sigma}$	: Matriks kovarian
$\boldsymbol{\rho}$	: Matriks korelasi
$\bar{Y}$	: Nilai *mean* dari variabel terikat
$\bar{X}_j$	: Nilai *mean* dari variabel bebas ke-j , $j = 1,2,\dots,k$
S_{x_j}	: *Standard deviation* dari X_j untuk $j = 1,2,\dots,k$
S_y	: *Standard deviation* dari variabel terikat
Z_{ij}	: Bentuk standar dari variabel X_{ij} , $i = 1,2,\dots,n$; $j = 1,2,\dots,k$
Y_i^*	: Bentuk standar dari variabel Y_i , dengan $i = 1,2,\dots,n$
$Fo\ (e_i)$	: Frekuensi relatif kumulatif, $i = 1,\dots,n$
$S_N(e_i)$	: Frekuensi teoritis (ekspektasi), $i = 1,\dots,n$
e_i	: Nilai residual ke- i, $i = 1,2,\dots,n$
dL	: Batas bawah tabel Durbin Watson
dU	: Batas atas tabel Durbin Watson
R_j^2	: Koefisien determinasi ke$-j$, $j = 1,2,\dots,k$
λ_j	: *Eigenvalue* ke$-j$, $j = 1,2,\dots,k$
$\boldsymbol{a}_j'$	: *Eigenvektor* ke-j, $j = 1,2,\dots,k$
w_0	: Konstanta *principal component regression*
$w_1,\dots,w_m$	: Parameter *principal component regression*
$K_1,\dots,K_m$	: Komponen utama
$Se(\hat{\beta}_j)$	: Standar *error* dari koefisien regresi $\hat{\beta}_j$, $j = 1,2,\dots,k$
SSR	: *Sum of Squares Regression*
SSE	: *Sum of Squares Error*

MSE	: *Mean Square Error*
SST	: *Sum of Squares Total*
MSR	: *Mean Square Regression*
n	: Banyaknya observasi
k	: Banyaknya variabel bebas
R_{adj}	: *Adjusted Coefficient Multiple Determination*
c_{HK}	: Konstanta bias *Hoerl and Kennard*
RSE	: *Residual Standard Error*

DAFTAR LAMPIRAN

ABSTRAK

Maulida, Rizka. 2022. **Perbandingan *Principal Component Regression* dan Regresi *Ridge* Pada Analisis Faktor-Faktor Indeks Pembangunan Manusia**. Skripsi. Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Fachrur Rozi, M.Si., (II) Mohammad Nafie Jauhari, M.Si.

Kata kunci: IPM, Regresi, Metode *Ordinary Least Square* (OLS), Multikolinieritas, *Principal Component Regression*, Regresi *Ridge*

Indeks Pembangunan Manusia (IPM) merupakan salah satu tolak ukur untuk menentukan pembangunan ekonomi. Provinsi Nusa Tenggara Timur masuk dalam 3 provinsi dengan indeks pembangunan manusia terendah. Pada model regresi linier berganda sering terjadi gejala multikolinieritas. Multikolinieritas adalah terjadinya korelasi linier antar sebagian atau seluruh variabel bebas yang menyebabkan estimasi parameter regresi tidak bisa ditaksir dengan akurat. Oleh karena itu, multikolinieritas perlu diatasi yaitu dengan metode *principal component regression* dan regresi *ridge*. Tujuan digunakannya metode tersebut adalah untuk mengatasi multikolinieritas yang tidak bisa dihilangkan dengan metode OLS. *Principal component regression* adalah regresi variabel terikat terhadap komponen utama yang tidak berkorelasi. Sedangkan, regresi *ridge* merupakan modifikasi dari metode OLS yaitu dengan menambahkan tetapan bias pada diagonal matriks $X'X$. Salah satu metode penentuan tetapan bias adalah *Hoerl & Kennard*. Data yang digunakan adalah data IPM Provinsi Nusa Tenggara Timur tahun 2021 yang dimodelkan menggunakan *principal component regression* dan regresi *ridge* menggunakan estimator parameter *Hoerl & Kennard* dengan melakukan langkah-langkah seperti uji asumsi klasik dan uji signifikansi parameter hingga diperoleh keakuratan model. Hasil analisis menggunakan metode OLS pada data IPM terdapat multikolinieritas yaitu variabel jumlah penduduk (X_1) dan jumlah penduduk usia 15 tahun yang bekerja (X_6) karena mempunyai nilai VIF lebih dari 10. Hasil analisis menggunakan *principal component regression* menghasilkan nilai VIF kurang dari 10. Sedangkan, hasil analisis menggunakan regresi *ridge* menggunakan estimator *Hoerl & Kennard* sebesar 0,2760 dan nilai VIF kurang dari 10. Kemudian, dari kedua metode dibandingkan berdasarkan R^2, R_{adj},dan RSE. Berdasarkan hasil analisis diperoleh bahwa nilai R^2,R_{adj} dari regresi *ridge* lebih tinggi dibandingkan *principal component regression* dan RSE regresi *ridge* lebih kecil daripada *principal component regression.* Kemudian diperoleh selisih R_{adj} sebesar 7,95 % yang lebih besar dari α. Oleh karena itu, dapat disimpulkan bahwa model regresi *ridge* lebih baik daripada *principal component regression* dalam mengatasi multikolinieritas pada data IPM di Nusa Tenggara Timur.

ABSTRACT

Maulida, Rizka. 2022. **On the Comparison of Principal Component Regression and Ridge Regression on Analysis of Human Development Index Factors**. Thesis. Mathematics Study Program, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Supervisor: (I) Fachrur Rozi, M.Si., (II) Mohammad Nafie Jauhari, M.Si.

Keywords: HDI, Regression, Ordinary Least Square (OLS) Method Multicollinearity, Principal Component Regression, Ridge Regression

The Human Development Index (HDI) is one of the benchmarks for determining economic development. East Nusa Tenggara Province is included in the three provinces with the lowest human development index. In the multiple linear regression model, symptoms of multicollinearity often occur. Multicollinearity is the occurrence of linear correlation between some or all of the independent variables which causes the estimated regression parameters cannot be estimated accurately. Therefore, multicollinearity needs to be overcome by using principal component regression and ridge regression methods. The purpose of using this method is to overcome multicollinearity that cannot be eliminated by the OLS method. Principal component regression is the regression of the dependent variable on the main component that is not correlated. Meanwhile, ridge regression is a modification of the OLS method by adding a bias constant to the diagonal of the $X'X$ matrix. One method of determining the bias constant is Hoerl & Kennard. The data used is the HDI data of East Nusa Tenggara Province in 2021 which is modeled using principal component regression and ridge regression using the Hoerl & Kennard parameter estimator by performing steps such as classical assumption test and parameter significance test to obtain model accuracy. The results of the analysis using the OLS method on HDI data have multicollinearity, namely the variable number of residents (X_1) and the number of people aged 15 years working (X_6) because they have a VIF value of more than 10. The results of the analysis using principal component regression resulted in a VIF value of less than 10. Meanwhile, the results of the analysis using ridge regression using the Hoerl & Kennard estimator were 0.2760 and the VIF value was less than 10. Then, the two methods were compared based on R^2,R_{adj}, and RSE. Based on the results of the analysis, it is found that the R^2,R_{adj} value of the ridge regression is higher than the principal component regression and the RSE of the ridge regression is smaller than the principal component regression. Then the difference in R_{adj} is 7.95% which is bigger than α. Therefore, it can be concluded that the ridge regression model is better than the principal component regression in overcoming multicollinearity in HDI data in East Nusa Tenggara.

مستخلص البحث

موليدة، رزقا. ٢٠٢٢. **المقارنة بين انحدار المكون الرئيسي (*Principal Component Regression*) وانحدار ريدج (*Ridge Regression*) في تحليل عوامل مؤشر التنمية البشرية**. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف الأول: فخر الرازي الماجستير، المشرف الثاني: محمد نافع جوهري الماجستير.

الكلمات المفتاحية: مؤشر التنمية البشرية، طريقة نهاية المربع العادية(*Ordinary Least Square Method*)، متعدد الخطية، انحدار المكون الرئيسي (*Principal Component Regression*)، انحدار ريدج (*Ridge Regression*).

مؤشر التنمية البشرية هو أحد المعايير لتعيين التنمية الاقتصادي. تقع نوسا الجنوب الشرقي من المحافظات الثلاثة الأقل في مؤشر التنمية البشرية. وفي نموذج الانحدار الخطي المزدوج، غالبًا ما تحدث عرض الخطية المتعددة. وأما الخطية المتعددة هي حدوث العلاقة خطي بين بعض أو كل المتغيرات المستقلة التي تتسبب إلى عدم إمكانية تقدير معيار الانحدار المقدرة بدقة. لذلك، يجب الحل لوجود الخطية المتعددة انحدار المكون الرئيسي وانحدار ريدج. والغرض من استخدام هذه الطريقة لحل الخطية المتعددة التي لا يتطيع أن تحلها بطريقة بطريقة نهاية المربع العادية (OLS). انحدار المكون الرئيسي هو انحدار المتغير التابع على المكون الرئيسي الذي غير المترابط. وأما انحدار ريدج هو تعديل لطريقة OLS وهو بإضافة ثابت التحيز إلى المصفوفة القطرية. وإحدى طرق لتعيين ثابت التحيز هو *Kennard&Hoerl*. والبيانات المستخدمة هي بيانات مؤشر التنمية البشرية لمحافظة نوسا تينجارا الشرقية عام 2021 والتي نمذجها باستخدام انحدار المكون الرئيسي وانحدار ريدج بمقدّر معيار *Kennard&Hoerl* بإجراء الخطوات مثل اختبار الافتراض الكلاسيكي واختبار معنوي المعيار للحصول على دقة النموذج. ودلت نتيجة التحليل بطريقة OLS لبيانات مؤشر التنمية البشرية إلى وجود علاقة الخطية المتعددة أي متغير عدد المسكان(X_1) وعدد السكان الذين يعملون في عمر 15 سنة (X_6) لأن لديهم قيمة VIF تزيد عن 10. ونتيجة التحليل باستخدام انحدار المكون الرئيسيينتج قيمة VIF أقل من 10. وفي الوقت نفسه، استخدمت نتيجة التحليل انحدار ريدج باستخدامقدر *Kennard&Hoerl* يبلغ إلى 0,2760 وقيمة VIF أقل من 10. ثم تقارن الطريقتين بناءً على R^2 وR_{adj} وRSE. وبناءً على نتائج التحليل، وجد أن قيمة R^2 وR_{adj} من انحدار ريج أعلى من انحدار المكون الرئيسيوأن قيمة RSE من انحدار ريدج أصغر من انحدار المكون الرئيسي. ثم حصلنا على فارق R_{adj} 7.95٪ وهو أكبر من α.لذلك، يستنتاج أن انحدار ريدج أفضل من انحدار المكون الرئيسيفي الحل على الخطية المتعددة على البيانات المؤشرة التنمية البشرية في محافظة نوسا الجنوب الشرقي .

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan teknik statistika untuk mencari pengaruh antara variabel bebas (X) terhadap variabel terikat (Y). Regresi linier antara variabel terikat dengan satu variabel bebas adalah regresi linier sederhana. Sedangkan regresi linier antara variabel terikat dengan dua atau lebih variabel bebas adalah regresi linier berganda (Kurtner & Nachtsheim, 2005).

Berdasarkan teori Gauss Markov, estimator yang baik untuk mengestimasi koefisien regresi yaitu menggunakan metode *Ordinary Least Square* (OLS). Karena estimator ini mempunyai sifat *Best Linear Unbiased Estimator* (BLUE) yaitu tidak bias dan memiliki varians yang minimum. Pada umumnya terdapat asumsi-asumsi yang wajib dipenuhi agar estimator regresi bisa ditaksir dengan metode OLS antara lain normalitas, tidak terjadi autokorelasi, homoskedastisitas, dan tidak terjadi multikolineraitas (Gujarati, 2009).

Dalam model regresi linier, pelanggaran asumsi yang banyak terjadi adalah multikolinieritas, yakni terdapat korelasi linier antar sebagian atau seluruh variabel bebas pada model regresi linier berganda (Gujarati, 2009). Akibatnya estimasi parameter regresi tidak bisa ditaksir dengan akurat sehingga parameter regresi menjadi bias. Beberapa metode untuk mengatasi multikolinieritas adalah regresi *ridge* dan *principal component regression* (Montgomery dan Peck, 1991).

Principal component regression membentuk hubungan antara variabel terikat dan komponen utama yang dipilih dari variabel bebas (UI-Saufie, dkk., 2011). Sedangkan regresi *ridge* merupakan modifikasi dari metode OLS dengan menambahkan tetapan bias untuk menstabilkan koefisien yang berguna untuk mengurangi varian (Mardikyan dan Cetin, 2008). Salah satu metode yang banyak digunakan untuk menentukan tetapan bias adalah metode *Hoerl and Kennard* .

Penelitian tentang pemilihan estimator parameter regresi *ridge* telah dilakukan peneliti sebelumnya. Fitrianto dan Yik (2014) mengevaluasi estimator regresi *ridge* menggunakan *Mean Square Error* (MSE) dengan simulasi Monte Carlo. Hasil dari penelitian yaitu parameter *Hoerl and Kennard* mempunyai MSE lebih kecil daripada OLS. Kemudian Solekakh, dkk. (2015), tentang data inflasi yang terdapat multikolinieritas menggunakan metode Hoerl, Kennard, dan Baldwin. Penelitian lain dilakukan oleh Hasriani (2016) yaitu perbandingan regresi *ridge* dan *principal component analysis* pada data hasil simulasi dan data kasus dari skripsi Nanang Pradipta. Dilihat dari MSE dan R^2 diperoleh hasil bahwa regresi *ridge* lebih baik dibandingkan *principal component analysis*.

Tirink, dkk. (2020), melakukan penelitian yaitu perbandingan regresi *ridge* dan metode OLS pada pengukuran tubuh anak. Dilihat dari MSE dan R^2 diperoleh bahwa metode OLS mempunyai nilai MSE yang lebih kecil dibandingkan metode regresi *ridge*, namun metode regresi *ridge* mempunyai nilai R^2 lebih tinggi daripada metode OLS. Dengan demikian, regresi *ridge* lebih baik dibandingkan OLS.

Kemudian Supriyadi, dkk. (2017), dalam penelitiannya yaitu "Perbandingan Metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) pada Kasus Pendapatan Anggaran Daerah Provinsi Jawa Tengah". Dengan melihat nilai R^2 dan nilai MSE diperoleh bahwa metode PLS lebih baik daripada metode PCR. Sriningsih, dkk. (2018), melakukan penelitian terhadap data impor beras di Provinsi Sulawesi Utara menggunakan *principal component regression* untuk mengatasi multikolinieritas. Penelitian mengenai Indeks Pembangunan Manusia di Jawa Tengah telah dilakukan oleh Faizia, dkk. (2019) menggunakan regresi komponen utama robust.

Pembangunan manusia adalah salah satu faktor penting dalam pertumbuhan ekonomi. Sumber daya manusia yang berkualitas mempunyai produktifitas tinggi sehingga dapat meningkatkan efisiensi ekonomi dan secara keseluruhan dapat mempengaruhi pertumbuhan ekonomi (Tjiptoherijanto, 1996). Indeks Pembangunan Manusia (IPM) berguna untuk membandingkan kinerja pembangunan orang lintas negara dan lintas wilayah (Kuncoro, 2006). Nusa Tenggara Timur (NTT) masuk dalam 3 besar provinsi dengan indeks pembangunan manusia terendah, dimana indeks pembangunan manusia menjadi tolak ukur pembangunan manusia. Indeks pembangunan manusia Provinsi Nusa Tenggara Timur adalah 65,19, terendah ketiga setelah Provinsi Papua dan Papua Barat. Di sisi lain, Provinsi Nusa Tenggara Timur juga mengalami penurunan indeks pembangunan pada tahun 2020, lebih rendah 0,06 poin dibandingkan tahun 2019. Sebagaimana firman Allah pada Al-Qur’an surat At-Taubah ayat 105 yaitu:

وَقُلِ اعْمَلُوْا فَسَيَرَى اللّٰهُ عَمَلَكُمْ وَرَسُوْلُهٗ وَالْمُؤْمِنُوْنَۗ وَسَتُرَدُّوْنَ اِلٰى عٰلِمِ الْغَيْبِ وَالشَّهَادَةِ فَيُنَبِّئُكُمْ بِمَا كُنْتُمْ

تَعْمَلُوْنَۚ ١٠٥

Artinya: ” *Katakanlah (Nabi Muhammad), “Bekerjalah! Maka, Allah, rasul*-*Nya*, *dan orang-orang mukmin akan melihat pekerjaanmu. Kamu akan dikembalikan kepada (Zat) yang mengetahui yang gaib dan yang nyata. Lalu, Dia akan memberitakan kepada kamu apa yang selama ini kamu kerjakan*.”

Dari ayat di atas dijelaskan bahwa umat manusia diperintahkan oleh Allah untuk selalu melakukan pekerjaan yang bermanfaat untuk diri sendiri dan orang lain. Karena semua amal akan dilihat oleh Allah, Rasul, dan para mukminin serta akan diperlihatkan oleh Allah di hari kiamat kelak. Lalu akan mendapat balasan sesuai amal perbuatannya ketika di muka bumi. Jika amal perbuatan yang baik akan mendapatkan pahala dan apabila perbuatannya jelek akan mendapat siksa (Muhammad Amin Suma, 2015).

Berdasarkan penjelasan latar belakang di atas penulis ingin menganalisis "Perbandingan *Principal Component Regression* dan Regresi *Ridge* pada Analisis Faktor-Faktor Indeks Pembangunan Manusia". Dari kedua metode tersebut akan dibandingkan nilai R^2 , R_{adj}, dan RSE.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini berdasarkan latar belakang di atas adalah bagaimana keakuratan model *principal component regression* dan regresi *ridge* dalam menganalisis faktor-faktor indeks pembangunan manusia?

1.3 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai berdasarkan rumusan masalah yang telah dijelaskan di atas adalah untuk menganalisis hasil keakuratan model *principal component regression* dan regresi *ridge* dalam menganalisis faktor-faktor indeks pembangunan manusia.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dalam penelitian ini antara lain:

1. Bagi penulis yaitu mampu menambah wawasan penulis terkait metode *principal component regression* dan regresi *ridge* pada analisis faktor-faktor indeks pembangunan manusia.
2. Bagi pembaca yaitu dapat memberikan tambahan ilmu dan wawasan tentang metode *principal component regression* dan regresi *ridge* untuk mengatasi multikolinieritas.
3. Bagi instansi yaitu diharapkan dapat memberikan masukan kepada pihak Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur untuk merencanakan program-program yang dapat meningkatkan indeks pembangunan manusia.

1.5 Batasan Masalah

Batasan masalah diperlukan agar tidak terjadi perluasan masalah yaitu:

1. Data yang digunakan merupakan data indeks pembangunan manusia provinsi Nusa Tenggara Timur tahun 2021 dengan diasumsikan variabel bebas terdapat multikolinieritas.
2. Estimator parameter yang digunakan dalam regresi *ridge* adalah *Hoerl and Kennard*.

BAB II

KAJIAN TEORI

2.1 Regresi Linier Berganda

Analisis regresi merupakan teknik statistika untuk mencari pengaruh antara variabel bebas (X) terhadap variabel terikat (Y). Regresi linier antara variabel terikat dengan satu variabel bebas adalah regresi linier sederhana. Sedangkan regresi linier antara variabel terikat dengan dua atau lebih variabel bebas adalah regresi linier berganda (Kurtner & Nachtsheim, 2005). Pertama kali Sir Francis Galton pada tahun 1877 memperkenalkan istilah regresi. Model regresi linier berganda dinyatakan dalam bentuk persamaan berikut (Montgomery, dkk., 2012).

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \tag{2.1}$$

dengan,

Y_i : Variabel terikat dengan $i = 1,2,\dots,n$

X_{ij} : Variabel bebas dengan $i = 1,2,\dots,n$; $j = 1,2,\dots,k$

β_j : Koefisien regresi ke-j dengan $j = 1,2,\dots,k$

ε_i : *Error* ke- i dengan $i = 1,2,\dots,n$

Jika dinyatakan dalam matriks sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.2}$$

Asumsi regresi linier berganda yang harus terpenuhi antara lain (Gujarati, 2009):

1. Nilai ekspektasi dari variabel ε_i sama dengan 0 atau

$$E(\varepsilon_i) = 0, \;\; i = 1,2,\ldots,n$$

2. Variansi konstan untuk setiap variabel ε_i dalam setiap pengamatan

$$Var\;(\varepsilon_i) = \sigma^2 \text{ untuk semua } i.$$

3. Tidak terjadi autokorelasi antar variabel ε_i

$$Cov\left(\varepsilon_i, \varepsilon_j\right) = 0, \quad i \neq j$$

4. Tidak terjadi multikolinieritas.

2.1.1 Metode *Ordinary Least Square* (OLS)

Metode *Ordinary Least Square* (OLS) adalah metode yang sering digunakan. Menurut Montgomery, dkk. (2012), estimasi parameter diperoleh dengan meminimumkan jumlah kuadrat *error*. Berdasarkan persamaan (2.1) maka diperoleh dalam bentuk matriks (2.2) kemudian persamaan tersebut dapat disederhanakan menjadi

$$\boldsymbol{Y} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{2.3}$$

Berdasarkan persamaan (2.3) diperoleh persamaan dugaan regresi adalah

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{2.4}$$

dan diperoleh jumlah kuadrat *error* yaitu

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 \tag{2.5}$$

Apabila diubah dalam bentuk matriks menjadi

$$S = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n] \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \tag{2.6}$$

Maka berdasarkan persamaan (2.4) dan (2.5) diperoleh

$$\begin{aligned} S &= \left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right)'\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \\ &= \left(\boldsymbol{Y}' - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\right)\left(\boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \\ &= \boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{Y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} \qquad (2.7) \\ &= \boldsymbol{Y}'\boldsymbol{Y} - (\boldsymbol{Y}'\boldsymbol{X}\widehat{\boldsymbol{\beta}})' - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} \\ &= \boldsymbol{Y}'\boldsymbol{Y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} \\ &= \boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{2}\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{Y} + \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} \end{aligned}$$

Dengan melakukan turunan parsial pertama S terhadap $\widehat{\boldsymbol{\beta}}$ untuk memperoleh nilai minimum dari persamaan tersebut diperoleh hasil yaitu

$$\frac{dS}{d\widehat{\boldsymbol{\beta}}} = 0 - 2\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} + \left(\widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\right)' \tag{2.8}$$

$$= -2\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{2}\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

Sehingga diperoleh

$$\widehat{\boldsymbol{\beta}}_{ols} = (\boldsymbol{X}'\boldsymbol{X})^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{Y} \tag{2.9}$$

dinamakan sebagai penaksir parameter $\boldsymbol{\beta}$ secara kuadrat terkecil (Aziz, 2010).

2.1.2 Pemusatan dan Penskalaan

Menurut Kurtner dan Nachtsheim (2005), pemusatan dan penskalaan merupakan bagian dari *standardized* variabel. Pemusatan dilakukan dengan menghilangkan $\hat{\beta}_0$ sehingga membuat perhitungan model regresi menjadi lebih sederhana. Sedangkan penskalaan dilakukan dengan mentransformasi variabel terikat dan variabel bebas dalam bentuk berikut:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_{X_j}} \tag{2.10}$$

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_Y} \tag{2.11}$$

$$S_{X_j} = \sqrt{\frac{\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2}{n-1}} \tag{2.12}$$

$$S_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}} \tag{2.13}$$

dengan,

Z_{ij} : Bentuk standar dari variabel X_{ij} , $i = 1,2,\ldots,n$; $j = 1,2,\ldots,k$

Y_i^* : Bentuk standar dari variabel Y_i , dengan $i = 1,2,\ldots,n$

$\bar{Y}$: Nilai *mean* dari variabel terikat

$\bar{X}_j$: Nilai *mean* dari variabel bebas ke-j, $j = 1,2,\ldots,k$

S_{X_j} : *Standard deviation* dari X_j untuk $j = 1,2,\ldots,k$

S_Y : *Standard deviation* dari variabel terikat

Sehingga diperoleh model regresi standar yaitu

$$Y_i^* = \beta_1^* Z_{1i} + \beta_2^* Z_{2i} + \beta_3^* Z_{3i} + \cdots + \beta_k^* Z_{ki} + \varepsilon_i^* \tag{2.14}$$

Hubungan antara parameter dari model persamaan (2.1) dan (2.14) adalah

$$\beta_j = \left(\frac{S_y}{S_{x_j}}\right)\beta_j^* , \qquad j = 1,2,\dots,k \tag{2.15}$$

$$\begin{aligned}\beta_0 &= \bar{Y} - \beta_1\bar{X}_1 - \beta_2\bar{X}_2 - \beta_3\bar{X}_3 - \cdots - \beta_k\bar{X}_k \\ &= \bar{Y} - \sum_{j=1}^{k}\beta_j\bar{X}_j\end{aligned} \tag{2.16}$$

2.1.3 Uji Asumsi Klasik

A. Uji Normalitas

Uji normalitas digunakan untuk memeriksa apakah residual berdistribusi normal atau tidak pada model regresi (Ghozali, 2016). Dalam penelitian ini digunakan uji *Kolmogorov-Smirnov*. Berikut ini merupakan pengujian hipotesis menggunakan *Kolmogorov-Smirnov*.

Hipotesis:

H_0 : *Error* data berdistribusi normal

H_1 : *Error* data tidak berdistribusi normal

Statistik Uji:

$$D = Max|Fo\ (e_i) - S_N(e_i)| \tag{2.17}$$

dengan,

$Fo\ (e_i)$: Frekuensi relatif kumulatif , $i = 1,\dots,n$

$S_N(e_i)$: Frekuensi teoritis (ekspektasi),$i = 1,\dots,n$

Keputusan:

Jika $p\ value > \alpha$ maka H_0 diterima.

B. Uji Heteroskedastisitas

Uji Heteroskedastisitas digunakan untuk mengetahui apakah residual dari model yang diteliti mempunyai varians yang konstan atau tidak. Model regresi yang baik adalah yang tidak terjadi heteroskedastisitas (Ghozali, 2016). Salah satu pengujian yang dilakukan yaitu uji Glejser. Uji Glejser dilakukan dengan melakukan regresi variabel bebas dan residual. Berikut adalah pengujian hipotesis menggunakan uji Glejser (Ghozali, 2016):

Hipotesis:

H_0 : Tidak ada heteroskedastisitas

H_1 : Ada heteroskedastisitas

Statistik uji:

$$t = \frac{\hat{\beta}_j}{Se(\hat{\beta}_j)} \tag{2.18}$$

dengan,

$\hat{\beta}_j$: Koefisien regresi ke- j , $j = 1,2, \dots, k$

$Se(\hat{\beta}_j)$: Standar *error* dari koefisien regresi $\hat{\beta}_j$, $j = 1,2, \dots, k$

Keputusan:

Jika $p\ value > \alpha$ maka H_0 diterima.

C. Uji Autokorelasi

Uji autokolerasi digunakan untuk memeriksa apakah terdapat hubungan dalam model regresi antara kesalahan pengganggu pada periode t dengan kesalahan pengganggu pada periode $t-1$. Autokorelasi terjadi karena observasi yang beruntun sepanjang waktu berkaitan satu sama lainnya (Ghozali, 2016). Pada penelitian ini menggunakan uji Durbin Watson (DW). Berikut merupakan pengujian hipotesis (Nurdin, 2018):

Hipotesis:

H_0 : Tidak terdapat autokorelasi

H_1 : Terdapat autokorelasi

Statistik Uji:

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \tag{2.19}$$

dengan,

e_i : Nilai residual ke-i, $i = 1,2, \dots, n$

e_{i-1} : Nilai residual satu periode sebelumnya

Keputusan:

Jika $d < dL$ atau $d > 4 - dL$ artinya terdapat autokorelasi.

Jika $dU < d < 4 - dU$ artinya tidak terdapat autokorelasi.

Jika $dL < d < dU$ atau $4 - dU < d < 4 - dL$ artinya tidak bisa ditarik kesimpulan.

dengan,

dL : Batas bawah tabel Durbin Watson

dU : Batas atas tabel Durbin Watson

Apabila keputusan yaitu tidak bisa ditarik kesimpulan maka dilakukan uji lain, yakni *run test*. *Run test* merupakan bagian dari *statistic* non-parametrik yang digunakan untuk menguji apakah antar residual terdapat korelasi (Ghozali, 2016). Adapun hipotesisnya sebagai berikut.

H_0: Residual acak (Tidak terjadi autokorelasi)

H_1: Residual tidak acak (Terjadi autokorelasi)

Statistik Uji:

Uji sampel besar (n_1 dan $n_2 > 20$)

$$Z = \frac{r - \bar{x}_r}{\sigma_r} \tag{2.20}$$

dengan,

$$\bar{x}_r = \frac{2n_1n_2}{n_1 + n_2} + 1$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}}$$

Keputusan:

Jika $p\ value < \alpha$ maka H_0 ditolak.

D. Uji Multikolinieritas

Tahun 1934 adalah kali pertama Ragnar Frisch memperkenalkan multikolinieritas yang artinya terdapat korelasi linier antar sebagian atau seluruh variabel bebas pada model regresi linier berganda (Gujarati, 2009). Multikolinieritas membuat varians dari parameter yang diestimasi lebih besar dari seharusnya sehingga tingkat akurasi dari estimasi akan menurun (Sukmono, 2014). Terdapat beberapa akibat apabila terjadi multikolinieritas tidak sempurna antara lain (Gujarati, 2009):

a. Standar *error* cenderung semakin meningkat dengan meningkatnya korelasi antar variabel bebas meskipun estimator OLS bisa ditentukan.

b. Dengan nilai standar *error* yang besar maka interval kepercayaan parameter populasi cenderung besar.

c. Apabila multikolinieritas tinggi maka kemungkinan untuk menerima hipotesis yang salah juga meningkat.

Terdapat beberapa cara untuk mendeteksi multikolinieritas antara lain:

a. Apabila R^2 tinggi, akan tetapi hanya sedikit variabel bebas yang signifikan (Gujarati, 2009).

b. Apabila koefisien korelasi mendekati satu atau lebih besar dari 0,75 (Marcus, dkk., 2012).

c. Apabila nilai *Varian Inflation Factor* (VIF) lebih besar dari 10 (Ryan, 1997).

Dalam penelitian ini untuk mendeteksi multikolinieritas menggunakan nilai (VIF) . Adapun pengujian hipotesis sebagai berikut (Ghozali, 2016):

Hipotesis:

H_0 : $VIF < 10$ (Tidak terdapat multikolinieritas)

H_1 : $VIF \geq 10$ (Terdapat multikolinieritas)

Statistik Uji:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2.21}$$

dengan,

R_j^2 : Koefisien determinasi ke$-j$, $j = 1,2, \dots, k$

Keputusan:

Jika $VIF \geq 10$ maka H_0 ditolak.

2.2 *Principal Component Analysis*

Terdapat dua tahap dalam *principal component regression* yaitu melakukan *principal component analysis* dan melakukan regresi variabel terikat dengan komponen utama (Notiragayu dan Nisa, 2008).

Principal component analysis merupakan metode statistika untuk mengubah sebagian besar variabel asli yang berkorelasi menjadi variabel baru yang lebih kecil dan tidak berkorelasi (Johnson dan Wichern, 2007). *Principal component analysis* bergantung pada matriks kovarian ($\mathbf{\Sigma}$) atau matriks korelasi ($\boldsymbol{\rho}$). Matriks kovarian digunakan jika variabel yang diamati mempunyai satuan yang sama. Akan tetapi, jika variabel yang diamati mempunyai satuan yang berbeda maka menggunakan matriks korelasi.

Principal component analysis yang didasarkan pada matriks kovarian ($\mathbf{\Sigma}$) lalu dibentuk variabel baru yaitu K adalah komponen utama yang merupakan kombinasi linier dari $X_1, X_2, \ldots, X_k$ dengan *eigenvalue* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ dan *eigenvector* $\boldsymbol{a}_{\boldsymbol{j}}' = [a_{1j}, a_{2j}, \ldots, a_{kj}]$, maka bentuk kombinasi linier sebagai berikut (Johnson dan Wichern, 2007):

$$\begin{aligned} K_1 &= \boldsymbol{a}_{\boldsymbol{1}}'\boldsymbol{X} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1k}X_k \\ K_2 &= \boldsymbol{a}_{\boldsymbol{2}}'\boldsymbol{X} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2k}X_k \\ &\vdots \\ K_k &= \boldsymbol{a}_{\boldsymbol{k}}'\boldsymbol{X} = a_{k1}X_1 + a_{k2}X_2 + \cdots + a_{kk}X_k \end{aligned} \tag{2.22}$$

dengan $Cov(K_j, K_1) = \boldsymbol{a}_{\boldsymbol{j}}'\,\mathbf{\Sigma}\boldsymbol{a}_{\boldsymbol{1}}$ dan $Var\ (K_j) = \boldsymbol{a}_{\boldsymbol{j}}'\,\mathbf{\Sigma}\boldsymbol{a}_{\boldsymbol{j}}$, $j = 1,2, \ldots, k$.

Eigenvalue (λ) dari matriks kovarian ($\mathbf{\Sigma}$) dihitung dengan syarat

$$|\mathbf{\Sigma} - \lambda \boldsymbol{I}| = 0 \tag{2.23}$$

Sedangkan, *eigenvector* $\boldsymbol{a_j'} = [a_{1j}, a_{2j}, \dots, a_{kj}]$ dari matriks kovarian ($\boldsymbol{\Sigma}$) dihitung menggunakan rumus yaitu:

$$(\boldsymbol{\Sigma} - \lambda \boldsymbol{I})\boldsymbol{a_j} = 0 \tag{2.24}$$

Keragaman total yang dapat dijelaskan oleh banyaknya komponen utama terpilih (m) yaitu:

$$\frac{\sum_{j=1}^{m} \lambda_j}{\sum_{j=1}^{k} \lambda_j} \times 100\% \quad , m \leq k \tag{2.25}$$

Komponen utama juga diperoleh dari variabel yang distandarkan adalah sebagai berikut:

$$Z_{ij} = \frac{(X_{ij} - \bar{X}_j)}{S_{x_j}}, \; j = 1,2, \dots, k \tag{2.26}$$

Komponen utama yang dibentuk sebagai kombinasi linier dari variabel yang distandarkan yaitu $Z_1, Z_2, \dots, Z_k$ dengan *eigenvalue* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ dan *eigenvector* $\boldsymbol{a_j'} = [a_{1j}, a_{2j}, \dots, a_{kj}]$, maka bentuk kombinasi linier sebagai berikut (Johnson dan Wichern, 2007):

$$\begin{aligned}
K_1 &= \boldsymbol{a_1'Z} = a_{11}Z_1 + a_{12}Z_2 + \cdots + a_{1k}Z_k \\
K_2 &= \boldsymbol{a_2'Z} = a_{21}Z_1 + a_{22}Z_2 + \cdots + a_{2k}Z_k \\
&\vdots \\
K_k &= \boldsymbol{a_k'Z} = a_{k1}Z_1 + a_{k2}Z_2 + \cdots + a_{kk}Z_k
\end{aligned} \tag{2.27}$$

dengan $E(\boldsymbol{Z}) = 0$ dan $Cov(\boldsymbol{Z}) = \boldsymbol{\rho}$. Matriks korelasi didefinisikan sebagai berikut:

$$\boldsymbol{\rho} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{kk} \end{bmatrix} \tag{2.28}$$

Nilai r_{ij} dihitung dengan rumus korelasi *Pearson* yaitu:

$$r_{ij} = \frac{n\sum Z_i Z_j - \sum Z_i \sum Z_j}{\sqrt{n\sum Z_i^2 - (\sum Z_i)^2}\sqrt{n\sum Z_i^2 - (\sum Z_i)^2}} \tag{2.29}$$

Eigenvalue (λ) dari matriks korelasi ($\boldsymbol{\rho}$) dihitung dengan syarat

$$|\boldsymbol{\rho} - \lambda \boldsymbol{I}| = \boldsymbol{0} \tag{2.30}$$

Sedangkan, *eigenvector* $\boldsymbol{a}_j' = [a_{1j}, a_{2j}, \dots, a_{kj}]$ dari matriks korelasi ($\boldsymbol{\rho}$) dihitung menggunakan rumus berikut:

$$(\boldsymbol{\rho} - \lambda \boldsymbol{I})\boldsymbol{a}_j = \boldsymbol{0} \tag{2.31}$$

Keragaman total yang dapat dijelaskan oleh komponen utama ke-j dari variabel yang distandarkan yaitu:

$$\frac{\lambda_j}{k} \times 100\%, \qquad j = 1,2,\dots,k \tag{2.32}$$

dengan,

k : Banyaknya variabel bebas

2.3 *Principal Component Regression*

Principal component regression merupakan salah satu metode untuk mengatasi multikolinieritas. *Principal component regression* adalah analisis regresi dari variabel terikat terhadap komponen utama yang tidak berkorelasi dengan masing-masing komponen utama merupakan kombinasi linear dari semua variabel bebas (Draper & Smith, 1992). Pemilihan komponen utama dapat dilakukan dengan cara yaitu:

a. Memilih komponen utama yang mempunyai keragaman kumulatif sebesar 75% (Morison, 1978).

b. Memilih *eigenvalue* yang lebih besar dari satu. Kriteria ini digunakan jika berdasakan matriks korelasi (Suryanto, 1988).

c. Jumiati (2018) merekomendasikan untuk melihat *scree plot*. *Scree plot* adalah plot antara *eigenvalue* λ_j dengan j. Untuk menentukan banyaknya komoponen utama dengan melihat patahan siku dari *scree plot*.

Tetapi, beberapa ahli merekomendasikan untuk memilih komponen utama yang mempunyai *eigenvalue* lebih dari satu. Jika *eigenvalue* kurang dari satu maka keragaman data yang dideskripsikan sangat kecil (Sriningsih, dkk., 2018).

Setelah diperoleh komponen utama melalui *principal component analysis*, selanjutnya adalah meregresikan komponen utama terpilih dengan variabel terikat.

2.3.1 *Principal component regression* berdasarkan matriks korelasi

Misalkan matriks $\boldsymbol{Q}$ adalah matriks orthogonal yang memenuhi persamaan $\boldsymbol{Q'Q} = \boldsymbol{QQ'} = \boldsymbol{I}.$ Maka persamaan regresi linier berganda menjadi *principal component regression* sebagai berikut:

$$\begin{aligned}\boldsymbol{Y} &= \boldsymbol{Z\beta}^* + \boldsymbol{\varepsilon} \\ &= \boldsymbol{ZI\beta}^* + \boldsymbol{\varepsilon} \\ &= \boldsymbol{Z}(\boldsymbol{Q'Q})\boldsymbol{\beta}^* + \boldsymbol{\varepsilon} \\ &= \boldsymbol{Z}(\boldsymbol{QQ'})\boldsymbol{\beta}^* + \boldsymbol{\varepsilon} \\ &= \boldsymbol{Kw} + \boldsymbol{\varepsilon}\end{aligned} \tag{2.33}$$

dengan $\boldsymbol{K} = \boldsymbol{ZQ}$ dan $\boldsymbol{w} = \boldsymbol{Q'\beta}^*$

Sehingga model *principal component regression* yang telah direduksi menjadi m komponen adalah

$$Y = w_0 + w_1 K_1 + w_2 K_2 + \cdots + w_m K_m + \varepsilon \tag{2.34}$$

dengan,

w_0 : Konstanta *principal component regression*

$w_1, w_{2,\ldots,} w_m$: Parameter *principal component regression*

$K_1, K_2, \ldots, K_m$: Komponen utama

m adalah banyaknya komponen utama terpilih dan $m \leq k$.

Komponen utama merupakan kombinasi linier dari variabel standar yaitu

$$\begin{aligned} K_1 &= a_{11} Z_1 + a_{21} Z_2 + \cdots + a_{k1} Z_k \\ K_2 &= a_{12} Z_1 + a_{22} Z_2 + \cdots + a_{k2} Z_k \\ &\vdots \\ K_m &= a_{1m} Z_1 + a_{2m} Z_2 + \cdots + a_{km} Z_k \end{aligned} \tag{2.35}$$

Jika persamaan (2.35) disubstitusi ke persamaan (2.34) maka diperoleh

$$\begin{aligned} Y &= w_0 + w_1(a_{11} Z_1 + a_{21} Z_2 + \cdots + a_{k1} Z_k) \\ &\quad + w_2(a_{12} Z_1 + a_{22} Z_2 + \cdots + a_{k2} Z_k) \\ &\quad + w_m(a_{1m} Z_1 + a_{2m} Z_2 + \cdots + a_{km} Z_k) + \varepsilon \\ &= w_0 + w_1 a_{11} Z_1 + w_1 a_{21} Z_2 + \cdots + w_1 a_{k1} Z_k + w_2 a_{12} Z_1 + \\ &\quad w_2 a_{22} Z_2 + \cdots + w_2 a_{k2} Z_k + w_m a_{1m} Z_1 + w_m a_{2m} Z_2 + \ldots + \\ &\quad w_m a_{km} Z_k + \varepsilon \\ &= w_0 + (w_1 a_{11} + w_2 a_{12} + \cdots + w_m a_{1m}) Z_1 + (w_1 a_{21} + w_2 a_{22} + \\ &\quad \cdots + w_m a_{2m}) Z_2 + (w_1 a_{k1} + w_2 a_{k2} + \cdots + w_m a_{km}) Z_k + \varepsilon \end{aligned} \tag{2.36}$$

Diperoleh persamaan regresi dugaan komponen utama sebagai berikut.

$$Y = b_0 + b_1 Z_1 + b_2 Z_2 + \cdots + b_k Z_k \tag{2.37}$$

dengan

$$b_0 = w_0$$

$$b_1 = w_1 a_{11} + w_2 a_{12} + \cdots + w_m a_{1m}$$

$$b_2 = w_1 a_{21} + w_2 a_{22} + \cdots + \ w_m a_{2m}$$

$$b_k = w_1 a_{k1} + w_2 a_{k2} + \cdots + w_m a_{km}$$

2.4 Regresi *Ridge*

Hoer dan R.W. Kennard pada tahun 1962 pertama kali memperkenalkan regresi *ridge*. Regresi *ridge* bertujuan mengatasi akibat korelasi yang tinggi antara beberapa variabel bebas dalam model regresi yang membuat matriks $\boldsymbol{X'X}$ hampir singular sehingga menghasilkan estimasi parameter model regresi yang tidak stabil (Draper dan Smith, 1992). Regresi *ridge* merupakan modifikasi dari metode OLS yang menghasilkan penduga bias dari koefisien regresi (Kutner, dkk., 2005). Didefinisikan $\boldsymbol{W} = \boldsymbol{XQ}$ dan $\boldsymbol{\alpha} = \boldsymbol{Q'\beta}$ dengan $\boldsymbol{Q}$ adalah matriks berukuran $k \times k$ sedemikian hingga $\boldsymbol{Q'Q} = \boldsymbol{QQ'} = \boldsymbol{I_k}$ dan $(\boldsymbol{XQ})'(\boldsymbol{XQ}) = \boldsymbol{\Lambda} =$ diag $(\lambda_1, \dots, \lambda_k)$, $\lambda_1, \dots, \lambda_k$ adalah *eigenvalue* dari matriks $\boldsymbol{X'X}$. Berdasarkan persamaan (2.3) maka diperoleh bentuk kanonik menurut Younker (2012) adalah

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \\ &= \boldsymbol{XI\beta} + \boldsymbol{\varepsilon} \\ &= \boldsymbol{X}(\boldsymbol{Q'Q})\boldsymbol{\beta} + \boldsymbol{\varepsilon} \end{aligned} \tag{2.38}$$

$$= \boldsymbol{X(QQ')\beta + \varepsilon}$$

$$= \boldsymbol{W\alpha +\ \varepsilon}$$

Diperoleh bentuk kanonik $\widehat{\boldsymbol{\alpha}}_{\boldsymbol{ols}}$ berdasarkan persamaan (2.9) yaitu:

$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{ols}} = \ (\boldsymbol{W'W})^{-\mathbf{1}}\boldsymbol{W'Y} \tag{2.39}$$

$$= ((\boldsymbol{XQ})'(\boldsymbol{XQ}))^{-\mathbf{1}}\boldsymbol{W'Y}$$

$$= \boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{W'Y}$$

Koefisien dugaan regresi *ridge* didefinisikan dalam persamaan berikut:

$$\boldsymbol{Y = X\beta + \varepsilon} \tag{2.40}$$

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \ \boldsymbol{X\beta}$$

$$\boldsymbol{\varepsilon'\varepsilon} = (\boldsymbol{Y} - \ \boldsymbol{X\beta})'(\boldsymbol{Y} - \ \boldsymbol{X\beta})$$

Dengan syarat pembatas:

$$\boldsymbol{\beta'\beta - k^2 = 0}$$

$$\boldsymbol{G} = (\boldsymbol{Y - X\beta})'(\boldsymbol{Y - X\beta}) + \boldsymbol{c}(\boldsymbol{\beta'\beta - k^2}) \tag{2.41}$$

Sehingga diperoleh

$$\boldsymbol{G} \ \ = (\boldsymbol{Y - X\beta})'(\boldsymbol{Y - X\beta}) + \boldsymbol{c}(\boldsymbol{\beta'\beta - k^2})$$

$$\boldsymbol{G} \ = \boldsymbol{Y'Y - Y'X\beta - \beta X'Y + \beta'X'X\beta +\ c}(\boldsymbol{\beta'\beta - k^2})$$

$$\boldsymbol{G = Y'Y - 2\beta'X'Y + \beta'X'X\beta +\ c}(\boldsymbol{\beta'\beta - k^2})$$

$$\boldsymbol{G = Y'Y - 2\beta'X'Y + \beta'X'X\beta +\ c\beta'\beta - ck^2}$$

$$\frac{\partial G}{\partial \boldsymbol{\beta}} = \boldsymbol{-2X'Y + 2X'X\beta + 2cI\beta = 0} \tag{2.42}$$

$\boldsymbol{\beta}$ yang memenuhi persamaan (2.42) merupakan penaksir beta dari regresi *ridge* sehingga diperoleh estimator regresi *ridge* dalam persamaan berikut:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{R}} = (\boldsymbol{X'X + cI})^{-\mathbf{1}}\boldsymbol{X'Y} \tag{2.43}$$

Berdasarkan persamaan (2.39) dan (2.43) maka diperoleh

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{\boldsymbol{R}} &= (\boldsymbol{W}'\boldsymbol{W} + c\boldsymbol{I})^{-1}\boldsymbol{W}'\boldsymbol{Y} \\ &= ((\boldsymbol{XQ})'(\boldsymbol{XQ}) + c\boldsymbol{I}))^{-1}\boldsymbol{W}'\boldsymbol{Y} \\ &= (\boldsymbol{\Lambda} + c\boldsymbol{I})^{-1}\boldsymbol{W}'\boldsymbol{Y} \end{aligned} \tag{2.44}$$

2.4.1 Estimator Parameter *Hoerl and Kennard*

Konstanta bias (c) didefinisikan sebagai berikut (Hoerl & Kennard, 1970).

$$c_{HK} = \frac{k\hat{\sigma}^2}{\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}}} \tag{2.45}$$

dengan,

c_{HK} : Konstanta bias *Hoerl and Kennard*

$\hat{\sigma}^2$: MSE dari metode OLS

$$\hat{\sigma}^2 = \frac{(\boldsymbol{Y} - \boldsymbol{W}\hat{\boldsymbol{\alpha}})'(\boldsymbol{Y} - \boldsymbol{W}\hat{\boldsymbol{\alpha}})}{n - k - 1} \tag{2.46}$$

$\hat{\boldsymbol{\alpha}}$: Estimasi parameter dari metode OLS

2.5 Uji Signifikansi Parameter

Setelah semua asumsi model regresi terpenuhi, langkah selanjutnya adalah menentukan apakah hipotesis diterima atau tidak yaitu dengan melakukan uji F dan uji t. Uji F dilakukan untuk mengetahui apakah variabel bebas berpengaruh terhadap variabel terikat secara simultan. Sedangkan, uji t dilakukan untuk mengetahui pengaruh dari setiap variabel bebas terhadap variabel terikat secara individu (Sulistyono & Sulistiyowati, 2017). Adapun, uji hipotesis menggunakan uji F sebagai berikut (Montgomery, dkk., 2012):

Hipotesis:

H_0 : $\beta^*_{R(1)}$=... $= \beta^*_{R(8)} = 0$ (Model regresi tidak signifikan)

H_1 : $\exists \beta^*_{R(j)} \neq 0,\ j = 1, \dots, 8$ (Model regresi signifikan)

Statistik Uji:

$$F_{hitung} = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE} \tag{2.47}$$

dengan,

SSR : *Sum of Squares Regression*

SSE : *Sum of Squares Error*

MSE : *Mean Square Error*

MSR : *Mean Square Regression*

n : Banyaknya observasi

Keputusan:

Jika $F_{hitung} > F_{tabel}$ atau $p - value < \alpha$ maka H_0 ditolak.

Adapun pengujian hipotesis menggunakan uji t sebagai berikut (Montgomery, dkk., 2012):

Hipotesis:

H_0: $\beta^*_{R(j)} = 0$ untuk $j = 1, \dots, 8$ (Koefisien regresi ke-j tidak signifikan)

H_1: $\beta^*_{R(j)} \neq 0$ untuk $j = 1, \dots, 8$ (Koefisien regresi ke-j signifikan)

Statistik Uji:

$$t = \frac{\hat{\beta}^*_{R(j)}}{Se(\hat{\beta}^*_{R(j)})} \tag{2.48}$$

Keputusan:

Jika $|t_{hitung}| > t_{tabel}$ atau $p - value < \alpha$ maka H_0 ditolak.

2.6 Kebaikan Model

1. Koefisien Determinasi

Untuk menentukan akurasi garis regresi yang terbentuk untuk mewakili kumpulan data yang diamati maka dilihat seberapa baik model yang diperoleh dapat menjelaskan kondisi sebenarnya yang disebut koefisien determinasi (R^2). Nilai R^2 merupakan ukuran yang menjelaskan sejauh mana kontribusi variabel bebas terhadap variabel terikat (Siagaan dan Sugiarto, 2006). Adapun rumus koefisen determinasi yaitu:

$$R^2 = 1 - \frac{SSE}{SST} \tag{2.49}$$

dengan,

SST : *Sum of Squares Total*

Menambahkan banyak variabel bebas ke dalam model maka selalu meningkatkan (R^2). Karena (R^2) sering dibuat besar dengan menambahkan variabel bebas maka disarankan agar ukuran ini dimodifikasi. Untuk mengoreksi R^2 digunakan koefisien determinasi ganda terkoreksi (*adjusted coefficient multiple determinati*on). Adapun Rumus R_{adj} sebagai berikut:

$$R_{adj} = 1 - \frac{SSE/(n-k)}{SST/(n-1)} \tag{2.50}$$

dengan,

R_{adj} : *Adjusted Coefficient Multiple Determination*

2. *Residual Standard Error* (RSE)

Residual Standard Error (RSE) merupakan estimasi standar deviasi dari ε. ε adalah selisih dari nilai yang diprediksi dengan nilai yang diamati. RSE didefinisikan sebagai berikut (James, dkk., 2013).

$$RSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-k-1}} = \sqrt{MSE} \tag{2.51}$$

2.7 Indeks Pembangunan Manusia

Indeks Pembangunan Manusia merupakan salah satu indikator untuk menentukan pembangunan ekonomi, mengukur tingkat kualitas fisik dan non fisik penduduk, yaitu kesehatan, tingkat pendidikan dan indikator angka ekonomi lainnya (Suliswanto, 2010). Jadi, manusia adalah kekayaan bangsa yang sesungguhnya. Pembangunan sumber daya manusia baik fisik maupun non fisik menyiratkan peningkatan kapasitas dasar penduduk. Pertumbuhan ekonomi dan pembangunan manusia saling bergantung dan saling berkontribusi satu sama lain (Dewi & Sutrisno, 2014).

1. Jumlah penduduk

   Penduduk sebagai penyedia tenaga kerja profesional, diperlukan untuk menghasilkan kegiatan ekonomi. Seiring dengan pertambahan penduduk mengakibatkan kebutuhan yang semakin bertambah dan semakin kompleks (Sukirno, 1985). Pertumbuhan penduduk dan faktor-faktor yang terkait dengan peningkatan jumlah penduduk aktif umumnya dianggap sebagai faktor positif untuk merangsang pertumbuhan ekonomi (Arsyad, 1997).

2. Tingkat Pengangguran Terbuka (TPT)

   Pengangguran terbuka adalah bagian dari angkatan bekerja yang tidak bekerja atau sedang mencari pekerjaan baik yang belum pernah bekerja maupun yang sudah bekerja atau mempersiapkan bisnis, mereka yang tidak mencari pekerjaan karena merasa tidak mungkin mencari pekerjaan dan mereka yang sudah memiliki pekerjaan dan mereka yang sudah memiliki pekerjaan tetapi belum memulai bekerja (Nurmainah, 2013). Tingkat pengangguran adalah persentase jumlah pengangguran terbuka terhadap jumlah angkatan kerja.

3. Penduduk miskin

Seseorang dikatakan miskin atau hidup dalam kemiskinan apabila pendapatan mereka atau akses mereka terhadap barang dan jasa relatif rendah dibanding orang lain dalam hal perekonomian. Selain itu, kemiskinan dapat dilihat sebagai tingkat absolut dari pendapatan atau standar hidup (Van den Berg, 2005).

4. Produk Dosmetik Regional Bruto (PDRB) per kapita

PDRB adalah nilai total atau harga pasar dari semua barang dan jasa akhir yang diproduksi dari suatu perekonomian selama kurun waktu tertentu (Nanga, 2001). PDRB per kapita sering digunakan sebagai tolak ukur rata-rata standar hidup masyarakat (Blanchard and Johnson, 2013). Semakin besar PDRB per kapita maka semakin makmur wilayah tersebut.

5. Angka Harapan Hidup (AHH)

Angka harapan hidup adalah tolak ukur yang penting dalam sebagai salah satu indikator untuk mengukur kualitas populasi (Supriatna, dkk., 2006). Umur panjang dan rendah dianggap dalam menggambarkan kemajuan sosial ekonomi perusahaan.

6. Jumlah Penduduk usia 15 tahun yang bekerja

Tenaga kerja adalah penduduk usia kerja (15-64) tahun atau jumlah seluruh penduduk dalam suatu negara yang dapat menghasilkan barang dan jasa apabila terdapat permintaan terhadap tenaga mereka dan apabila mereka ingin berpartisipasi dalam aktivitas tersebut (Mulyadi, 2003).

7. Fasilitas Kesehatan

Derajat kesehatan masyarakat di negara dipengaruhi oleh adanya fasilitas kesehatan. Undang-Undang Nomor 36 Tahun 2009 tentang kesehatan menyatakan bahwa fasilitas pelayanan kesehatan adalah alat yang digunakan untuk menyelenggarakan upaya pelayanan kesehatan, baik promotif, preventif, kuratif maupun rehabilitatif yang dilakukan oleh pemerintah, pemerintah daerah, dan masyarakat.

8. Rata-rata lama sekolah

Pendidikan adalah salah satu bentuk investasi bagi setiap individu. Semakin tinggi tingkat pendidikan maka semakin tinggi kesejahteraan individu yang juga akan mempengaruhi kesejahteraan ekonomi jangka panjang suatu negara (Mankiw, 2012)

2.8 Perintah Bekerja dalam Islam

Menurut Asyraf Hj Ab Rahman Istilah kerja dalam Islam bukan semata-mata merujuk dalam mencari rezeki untuk menghidupi diri dan keluarga dengan menghabiskan waktu dari pagi hingga sore, tetapi mencakup segala pekerjaan yang mempunyai unsur kebaikan dan keberkahan bagi diri, keluarga, masyarakat, dan juga negara (Noor, 2012). Sebagaimana dalam QS. Al-Mulk ayat 15 yaitu:

هُوَ الَّذِيْ جَعَلَ لَكُمُ الْاَرْضَ ذَلُوْلًا فَامْشُوْا فِيْ مَنَاكِبِهَا وَكُلُوْا مِنْ رِّزْقِهٖۗ وَاِلَيْهِ النُّشُوْرُ ١٥

Artinya: *"Dialah yang menjadikan bumi untuk kamu dalam keadaan mudah dimanfaatkan. Maka, jelajahilah segala penjurunya dan makanlah sebagian dari rezeki-Nya. Hanya kepada-Nya kamu (kembali setelah) dibangkitkan."*

Berdasarkan ayat di atas dijelaskan bahwa ajakan manusia untuk memanfaatkan bumi dengan segala aktivitas baik bertani, berniaga, dan lain-lain. Diperintahkan untuk menggunakannya yang berguna demi kenyamanan hidup di dunia dengan tidak melupakan generasi sesudahnya. Karena tidak mungkin manusia bisa menghabiskan mengingat banyak sekali rezeki-Nya yang melebihi kebutuhan manusia itu sendiri. Dan dalam ayat terakhir ditutup dengan peringatan yang akan ada pertanggungjawaban atas semua amalan manusia di hadapan Allah (Shihab, 2016).

Oleh karena itu, dalam islam bekerja mempunyai posisi yang begitu mulia. Islam sangat menghargai seseorang yang bekerja dengan tangannya sendiri untuk memenuhi kebutuhan hidupnya (Alydrus, 2009). Saat seseorang merasa lelah setelah pulang bekerja maka Allah SWT mengampuni dosa-dosanya saat itu juga. Selain itu, seseorang yang bekerja dan berusaha untuk memperoleh penghasilan dengan tangannya sendiri secara halal yang bertujuan untuk memenuhi kebutuhannya sendiri atau kebutuhan orang yang menjadi tanggungannya maka seseorang ini dikategorikan sebagai *jihad fi sabiilillah*.

2.9 Kajian Topik dengan Teori Pendukung

Analisis regresi digunakan untuk mencari pengaruh antara variabel bebas terhadap variabel terikat. Untuk mengestimasi koefisien regresi digunakan metode OLS. Setelah dilakukan estimasi koefisien regresi, maka dilakukan pengujian asumsi klasik. Salah satunya yaitu uji multikolinieritas. Apabila dalam regresi linier berganda terdapat multikolinieritas maka sulit mendapatkan estimasi yang akurat. Oleh karena itu, digunakan metode untuk mengatasi multikolinieritas yaitu *principal component regression* dan regresi *ridge*.

Ada dua cara membentuk *principal component regression* berdasarkan *principal component analysis* yaitu melalui matriks kovarian dan matriks korelasi. Langkah pertama, melakukan standarisasi terhadap variabel bebas (2.26) dikarenakan data yang akan diteliti tidak mempunyai satuan yang sama. Selanjutnya, mencari *eigenvalue* (2.30) dan *eigenvector* (2.31) dari matriks korelasi. Pemilihan komponen utama ditentukan dengan melihat *eigenvalue* yang lebih besar dari satu. Kemudian, menghitung skor komponen utama. Skor komponen utama yang diperoleh lalu diregresikan dengan variabel terikat (2.34). Lalu mengembalikan persamaan regresi dalam bentuk variabel standar (2.37). Setelah itu, dilakukan uji signifikansi parameter yaitu uji t pada hasil *principal component regression*. Langkah terakhir, mengembalikan persamaan dalam bentuk variabel asli.

Sedangkan tahapan regresi *ridge* sebagai berikut. Pertama, melakukan transformasi data menggunakan pemusatan dan penskalaan (2.10) dan (2.11). Kedua, menentukan tetapan bias menggunakan metode *Hoerl and Kennard* (2.45). Setelah diperoleh tetapan bias, maka dilakukan estimasi parameter regresi *ridge* dengan metode *Hoerl and Kennard.* Selanjutnya, dilakukan uji signifikansi parameter.

Langkah terakhir adalah membandingkan kedua metode berdasarkan nilai R^2 (2.49), R_{adj} (2.50) dan RSE (2.51).

BAB III

METODE PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Penelitian deskriptif kuantitatif adalah penelitian yang bertujuan untuk membuat deskriptif tentang keadaan yang menggunakan angka mulai dari pengumpulan sampai dengan penafsiran hasil.

3.2 Data dan Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Nusa Tenggara Timur 2021. Data yang digunakan dalam penelitian ini adalah data indeks pembangunan manusia tahun 2021. Variabel dalam penelitian ini menggunakan indeks pembangunan manusia sebagai variabel terikat dan variabel bebas sebanyak delapan yang diduga mempengaruhi variabel terikat. Berikut adalah variabel yang digunakan.

Tabel 3.1 Variabel Penelitian

Simbol	Variabel	Satuan
Y	Indeks Pembangunan Manusia (IPM)	%
X_1	Jumlah Penduduk	Ribu
X_2	Tingkat Pengangguran Tebuka (TPT)	%
X_3	Penduduk miskin	%
X_4	Produk Domestik Regional Bruto (PDRB) per kapita	Rp
X_5	Angka Harapan Hidup (AHH)	Tahun
X_6	Penduduk usia 15 tahun yang bekerja	Jumlah
X_7	Fasilitas kesehatan	Jumlah
X_8	Rata-rata lama sekolah	Tahun

3.3 Instrumen Penelitian

Instrumen penelitian yang digunakan adalah dokumentasi. Dalam hal ini, penulis mempelajari dokumen tertulis yang sudah dipublikasi yaitu dalam buku " Provinsi Nusa Tenggara Timur Dalam Angka Tahun 2021". Sedangkan instrumen penelitian untuk mengolah data digunakan *software* yaitu SPSS, Minitab, dan Matlab.

3.4 Teknik Analisis Data

Teknik analisis data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Implementasi model regresi linier berganda pada data indeks pembangunan manusia.
   a. Melakukan deskripsi pada data indeks pembangunan manusia.
   b. Estimasi parameter menggunakan metode OLS yang diduga terdapat multikolinieritas pada model regresi.
   c. Melakukan uji asumsi klasik antara lain:
      i. Uji Normalitas
      ii. Uji Heteroskedastisitas
      iii. Uji Autokorelasi
      iv. Uji Multikolinieritas

   Jika terdapat multikolinieritas maka dilakukan penanganan menggunakan metode yaitu *principal component regression* dan regresi *ridge*.
2. Implementasi *principal component regression*, dengan tahapan sebagai berikut:

a. Melakukan standarisasi terhadap variabel bebas.

b. Menghitung *eigenvalue* dari matriks korelasi.

c. Menghitung skor komponen utama .

d. Melakukan regresi dengan komponen utama terpilih yaitu meregresikan variabel terikat terhadap skor komponen utama.

e. Mengembalikan persamaan regresi dalam bentuk variabel standar.

f. Mengembalikan persamaan regresi dalam bentuk variabel awal.

3. Implementasi regresi *ridge* dengan tahapan sebagai berikut:

a. Melakukan transformasi menggunakan pemusatan dan penskalaan.

b. Menentukan konstanta bias dengan metode *Hoerl and Kennard*.

c. Melakukan estimasi parameter model regresi *ridge* menggunakan estimator parameter *Hoerl and Kennard*.

d. Melakukan uji signifikansi parameter.

4. Membandingkan model *principal component regression* dan regresi *ridge* berdasarkan nilai R^2, R_{adj}, dan RSE.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Implementasi Model Regresi Linier Berganda pada Data Indeks Pembangunan Manusia

4.1.1 Deskripsi Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi NTT tahun 2021 dengan delapan variabel bebas yang diduga berpengaruh terhadap IPM. Adapun deskripsi data sebagai berikut:

Tabel 4.1 Deskripsi Statistik

Variabel	***Mean***	**Standar Deviasi**
Y	64,10	4,23
X_1	242,07	101,38
X_2	4,15	1,88
X_3	21,61	7,26
X_4	11704,95	6082,67
X_5	66,15	2,17
X_6	123907,05	53725,69
X_7	82,05	35,43
X_8	7,52	1,10

Adapun penjelasan pada Tabel 4.1 sebagai berikut:

a. Variabel Y, yaitu IPM mempunyai nilai *mean* sebesar 64,10 dan nilai standar deviasi sebesar 4,23.

b. Variabel X_1, yaitu jumlah penduduk mempunyai nilai *mean* sebesar 242,07 dan nilai standar deviasi sebesar 101,38.

c. Variabel X_2, yaitu TPT mempunyai nilai *mean* sebesar 4,15 dan nilai standar deviasi sebesar 1,88.

d. Variabel X_3, yaitu persentase penduduk miskin mempunyai nilai *mean* sebesar 21,61 dan nilai standar deviasi sebesar 7,26.
e. Variabel X_4, yaitu PDRB mempunyai nilai *mean* sebesar 11704,95 dan nilai standar deviasi sebesar 6082,67.
f. Variabel X_5, yaitu AHH mempunyai nilai *mean* sebesar 66,15 dan nilai standar deviasi sebesar 2,17.
g. Variabel X_6, yaitu penduduk usia 15 tahun yang bekerja mempunyai nilai *mean* sebesar 123907,05 dan nilai standar deviasi sebesar 53725,69.
h. Variabel X_7, yaitu jumlah fasilitas kesehatan mempunyai nilai *mean* sebesar 82,05 dan nilai standar deviasi sebesar 35,43.
i. Variabel X_8, yaitu rata-rata lama sekolah mempunyai nilai *mean* sebesar 7,52 dan nilai standar deviasi sebesar 1,10.

Nilai *mean* dan standar deviasi dari masing-masing variabel pada Tabel 4.1 akan digunakan dalam perhitungan standarisasi data dan perhitungan metode pemusatan dan penskalaan. Berdasarkan Tabel 4.1 diketahui bahwa satuan variabel yang digunakan berbeda sehingga rentang data yang diperoleh cukup jauh sehingga dilakukan standarisasi.

4.1.2 Estimasi Metode OLS

Menggunakan persamaan (2.9) diperoleh estimasi parameter menggunakan metode OLS pada Tabel 4.2.

Tabel 4.2 Estimasi Parameter Metode OLS

Variabel	**Estimasi** ($\hat{\boldsymbol{\beta}}_j$)
Konstanta	0,000
Z_1	0,014
Z_2	-0,082
Z_3	-0,092
Z_4	0,536
Z_5	0,369
Z_6	-0,091
Z_7	0,161
Z_8	0,322

Setelah diperoleh estimasi parameter β, selanjutnya dilakukan uji F untuk mengetahui signifikansi model regresi. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : \beta_0 = \beta_1 = ... = \beta_8 = 0$ (Model regresi tidak signifikan)

$H_1 : \exists \beta_j \neq 0,\ j = 0,1,...,8$ (Model regresi signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}$. Adapun hasil uji F menggunakan persamaan (2.47) adalah 25,297, sedangkan dengan $\alpha = 5\%$ diperoleh $F_{(0,05;8;13)} = 2,77$. Karena $F_{hitung} > F_{tabel}$ maka H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa parameter β berpengaruh secara signifikan.

Pada model regresi yang terbentuk berikutnya dilakukan uji t yaitu untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : \beta_j = 0\ , j = 0,1,...,8$ (Koefisien regresi ke-j tidak signifikan)

$H_1 : \beta_j \neq 0,\ j = 0,1,...,8$ (Koefisien regresi ke- j signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $|t_{hitung}| > t_{tabel}$. Hasil uji t menggunakan persamaan (2.48) dapat dilihat pada Tabel 4.3.

Tabel 4.3 Uji t Metode OLS

Variabel	t_{hitung}	$\lvert t_{hitung}\rvert$	t_{tabel}	**Keputusan**
Konstanta	0,000	0,000	2,16037	Tidak Signifikan
Z_1	0,022	0,022		Tidak Signifikan
Z_2	-0,663	0,663		Tidak Signifikan
Z_3	-0,932	0,932		Tidak Signifikan
Z_4	2,756	2,756		Signifikan
Z_5	4,580	4,580		Signifikan
Z_6	-0,149	0,149		Tidak Signifikan
Z_7	1,472	1,472		Tidak Signifikan
Z_8	2,201	2,201		Signifikan

Pada Tabel 4.3 diketahui bahwa variabel Z_4 yaitu PDRB, variabel Z_5 yaitu AHH, variabel Z_8 yaitu rata-rata lama sekolah berpengaruh signifikan terhadap IPM. Sedangkan, variabel Z_1 yaitu jumlah penduduk, variabel Z_2 yaitu TPT, variabel Z_3 yaitu persentase penduduk miskin, variabel Z_6 yaitu penduduk usia 15 tahun yang bekerja, dan variabel Z_7 yaitu jumlah fasilitas kesehatan tidak berpengaruh signifikan terhadap IPM.

4.1.3 Uji Asumsi Klasik

1. Uji Normalitas

Pengujian normalitas pada penulisan ini menggunakan uji *Kolmogorov Smirnov*. Adapun hipotesis yang digunakan yaitu:

H_0: *Error* data berdistribusi normal

H_1: *Error* data tidak berdistribusi normal

Dengan menggunakan taraf signifikan $\alpha = 5\%$ dan kriteria uji yaitu jika $p\ value > \alpha$ maka H_0 diterima. Hasil dari uji *Kolmogorov Smirnov* dengan persamaan (2.17) adalah 0,104, sedangkan $p\ value$ diperoleh sebesar 0,200.

Hal ini menunjukkan bahwa $p\,value > \alpha$ maka H_0 diterima artinya *error* data berdistribusi normal. Dengan demikian, asumsi normalitas terpenuhi.

2. Uji Heteroskedastisitas

Pengujian heteroskedastisitas pada penelitian ini menggunakan uji Glejser. Adapun hasil pengujian sebagai berikut:

Hipotesis:

H_0: Tidak ada heteroskedastisitas

H_1: Ada heteroskedastisitas

Dengan menggunakan taraf signifikan $\alpha = 5\%$ dan kriteria uji yaitu jika $p\,value > \alpha$ maka H_0 diterima. Hasil dari uji Glejser menggunakan persamaan (2.18) dapat dilihat pada tabel berikut.

Tabel 4.4 Uji Glejser pada Uji Heteroskedastisitas

Variabel	t_{hitung}	$p\,value$	**Keputusan**
Z_1	-0,526	0,608	Tidak Signifikan
Z_2	-1,777	0,099	Tidak Signifikan
Z_3	-0,065	0,949	Tidak Signifikan
Z_4	0,973	0,349	Tidak Signifikan
Z_5	1,242	0,236	Tidak Signifikan
Z_6	0,383	0,708	Tidak Signifikan
Z_7	0,108	0,916	Tidak Signifikan
Z_8	0,058	0,955	Tidak Signifikan

Tabel 4.4 diketahui bahwa pada hasil uji Glejser menunjukkan bahwa semua variabel bebas memiliki $p\,value > \alpha$ maka dapat disimpulkan bahwa H_0 diterima yang artinya tidak ada heteroskedastisitas.

3. Uji Autokorelasi

Pengujian autokorelasi pada penulisan ini menggunakan uji Durbin Watson. Adapun hipotesis sebagai berikut:

Hipotesis:

H_0: Tidak ada autokorelasi

H_1: Ada autokorelasi

Hasil pengujian Durbin Watson (DW) menggunakan persamaan (2.19) disajikan pada Tabel 4.5.

Tabel 4.5 Hasil Uji Autokorelasi menggunakan DW

Durbin Watson	
d	1,885
dL	0,5884
dU	2,4072

Berdasarkan Tabel 4.5 diketahui bahwa $dL < d < dU$ artinya tidak bisa ditarik kesimpulan. Karena tidak dapat ditarik kesimpulan maka dilakukan uji lain yaitu *run test*. Adapun hipotesisnya sebagai berikut:

H_0: Tidak ada autokorelasi

H_1: Ada autokorelasi

Dengan menggunakan taraf signifikan $\alpha = 5\%$ dan kriteria uji yaitu jika $p\ value > \alpha$ maka H_0 diterima. Hasil dari uji *run test* menggunakan persamaan (2.20) adalah 0,000 dengan $p\ value$ sebesar 1,000 artinya $p\ value > 0{,}05$ sehingga terima H_0 yaitu tidak ada autokorelasi.

4. Uji Multikolinieritas

Pengujian multikolinieritas pada penulisan ini menggunakan nilai VIF menggunakan persamaan (2.21). Apabila nilai VIF ≥ 10, maka terdapat multikolinieritas. Adapun hipotesis yang digunakan sebagai berikut:

H_0: $VIF < 10$ (Tidak ada multikolinieritas)

H_1: $VIF \geq 10$ (Terdapat multikolinieritas)

Tabel 4.6 Nilai VIF pada Uji Multikolinieritas

Variabel Bebas	VIF	Keputusan
Z_1	85,076	Ada Multikolinieritas
Z_2	3,331	Tidak Ada Multikolinieritas
Z_3	2,106	Tidak Ada Multikolinieritas
Z_4	8,158	Tidak Ada Multikolinieritas
Z_5	1,401	Tidak Ada Multikolinieritas
Z_6	80,180	Ada Multikolinieritas
Z_7	2,565	Tidak Ada Multikolinieritas
Z_8	4,615	Tidak Ada Multikolinieritas

Pada Tabel 4.6 diketahui bahwa variabel jumlah penduduk (Z_1) dan jumlah penduduk usia 15 tahun yang bekerja (Z_6) terdapat multikolinieritas di mana nilai VIF ≥ 10, maka dapat dikatakan model regresi memiliki multikolinieritas. Untuk mengatasi masalah tersebut maka digunakan metode *principal component regression* dan regresi *ridge*.

4.2 *Principal Component Regression*

4.2.1 Standarisasi data

Dalam penelitian ini data yang digunakan memiliki satuan yang berbeda maka langkah pertama yaitu mentransformasi variabel bebas ke dalam variabel standar menggunakan rumus (2.26). Adapun hasil yang diperoleh terdapat pada lampiran 2.

4.2.2 Menghitung *eigenvalue* dari matriks korelasi

Dalam penelitian ini data yang digunakan mempunyai satuan yang tidak sama sehingga digunakan matriks korelasi. Adapun hasil matriks korelasi dapat dilihat pada Tabel 4.7.

Tabel 4.7 Matriks Korelasi

	Z(X1)	Z(X2)	Z(X3)	Z(X4)	Z(X5)	Z(X6)	Z(X7)	Z(X8)
Z(X1)	1	0,232	-0,309	0,424	0,293	0,978	0,629	0,318
Z(X2)	0,232	1	-0,437	0,815	0,239	0,099	0,053	0,739
Z(X3)	-0,309	-0,437	1	-0,446	-0,175	-0,201	-0,055	-0,587
Z(X4)	0,424	0,815	-0,446	1	0,266	0,274	0,162	0,826
Z(X5)	0,293	0,239	-0,175	0,266	1	0,254	-0,136	0,2
Z(X6)	0,978	0,099	-0,201	0,274	0,254	1	0,678	0,202
Z(X7)	0,629	0,053	-0,055	0,162	-0,136	0,678	1	0,173
Z(X8)	0,318	0,739	-0,587	0,826	0,2	0,202	0,173	1

Untuk menentukan jumlah komponen utama yang terbentuk adalah dengan melihat *eigenvalue* dari matriks korelasi yang lebih besar dari satu. Adapun *eigenvalue* diperoleh menggunakan persamaan (2.30) yang disajikan pada Tabel 4.8.

Tabel 4.8 *Eigenvalue* dari Matriks Korelasi

Variabel	***Eigenvalue***	***Cumulative***
Z_1	3,6372	0,455
Z_2	2,0095	0,706
Z_3	1,0285	0,834
Z_4	0,6706	0,918
Z_5	0,3003	0,956
Z_6	0,2293	0,984
Z_7	0,1187	0,999
Z_8	0,0059	1,000

Pada Tabel 4.8 dapat diketahui bahwa variabel yang memiliki *eigenvalue* lebih dari satu adalah Z_1, Z_2, Z_3 sebanyak tiga. Variabel Z_1 mempunyai *eigenvalue* sebesar 3,6372 artinya keragaman ini dapat menjelaskan 45,5 % dari total *cumulative*. Variabel Z_2 mempunyai *eigenvalue* sebesar 2,0095 artinya keragaman ini dapat menjelaskan 70,6% dari total *cumulative.* Sedangkan, variabel Z_3 mempunyai *eigenvalue* sebesar 1,0285 artinya keragaman ini dapat menjelaskan 83,4% dari total *cumulative*.

4.2.3 Menghitung Skor Komponen Utama

Langkah selanjutnya adalah mencari skor komponen yang terbentuk untuk membentuk persamaan *principal component regression*. Skor komponen ini akan menggantikan variabel standar dengan variabel baru. Berikut adalah skor komponen yang dapat dilihat pada Tabel 4.9.

Tabel 4.9 Nilai Skor Komponen Koefisien Matriks

Variabel	K_1	K_2	K_3
Z_1	0,402	0,417	0,124
Z_2	0,379	-0,366	-0,105
Z_3	-0,323	0,206	0,046
Z_4	0,440	-0,261	-0,092
Z_5	0,198	-0,068	0,876
Z_6	0,348	0,503	0,126
Z_7	0,244	0,483	-0,380
Z_8	0,422	-0,300	-0,188

Berdasarkan Tabel 4.9, diperoleh persamaan *principal component regression* sebagai berikut.

$$\begin{aligned} K_1 &= 0{,}402\ Z_1 + 0{,}379\ Z_2 - 0{,}323\ Z_3 + 0{,}440\ Z_4 + 0{,}198\ Z_5 \\ &\quad + 0{,}348\ Z_6 + 0{,}244\ Z_7 + 0{,}422\ Z_8 \\ K_2 &= 0{,}417\ Z_1 - 0{,}366\ Z_2 + 0{,}206\ Z_3 - 0{,}261\ Z_4 - 0{,}068\ Z_5 \\ &\quad + 0{,}503\ Z_6 + 0{,}483\ Z_7 - 0{,}300\ Z_8 \\ K_3 &= 0{,}124\ Z_1 - 0{,}105\ Z_2 + 0{,}046\ Z_3 - 0{,}092\ Z_4 + 0{,}876\ Z_5 \\ &\quad + 0{,}126\ Z_6 - 0{,}380\ Z_7 - 0{,}188\ Z_8 \end{aligned} \tag{4.1}$$

Tabel variabel baru komponen utama yang terbentuk dengan bantuan *software* Minitab terdapat pada lampiran 3. Variabel tersebut menunjukkan bahwa tiga komponen adalah jumlah optimal untuk mereduksi ke sepuluh variabel bebas.

4.2.4 Meregresikan Variabel Terikat Terhadap Komponen Utama

Berdasarkan lampiran 3, selanjutnya K_1, K_2, dan K_3 atau variabel baru komponen utama sudah bisa digunakan untuk analisis regresi sehingga diperoleh model *principal component regression* sebagai berikut:

$$Y = 64{,}104 + 1{,}931\ K_1 - 0{,}902\ K_2 + 0{,}623\ K_3 \qquad (4.2)$$

Setelah diperoleh model pada persamaan (4.2) selanjutnya, dilakukan uji F untuk mengetahui signifikansi model regresi. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : w_0 = w_1 = \ldots = w_3 = 0$ (Model regresi tidak signifikan)

$H_1 : \exists w_j \neq 0,\ j = 0, \ldots, 3$ (Model regresi signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}$. Adapun hasil uji F menggunakan persamaan (2.47) adalah 40,624, sedangkan dengan $\alpha = 5\%$ diperoleh $F_{(0{,}05;3;18)} = 3{,}16$. Karena $F_{hitung} > F_{tabel}$ maka H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa model regresi berpengaruh secara signifikan.

Uji signifikansi parameter selanjutnya yaitu uji t untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : w_j = 0$, $j = 0, \ldots, 3$ (Koefisien regresi ke- j tidak signifikan)

$H_1 : w_j \neq 0, j = 0, \ldots, 3$ (Koefisien regresi ke- j signifikan)

Kriteria uji yaitu jika $|t_{hitung}| > t_{tabel}$ maka H_0 ditolak. Nilai t_{hitung} diperoleh menggunakan persamaan (2.48). Adapun hasilnya disajikan pada Tabel 4.10.

Tabel 4.10 Uji t Metode *Principal Component Regression*

Variabel	t_{hitung}	$\lvert t_{hitung} \rvert$	Keputusan	VIF
Konstanta	183,367	183,367	Signifikan	
K_1	10,294	10,294	Signifikan	1,000
K_2	-3,574	3,574	Signifikan	1,000
K_3	1,767	1,767	Tidak Signifikan	1,000

Sedangkan, nilai t_{tabel} diperoleh sebesar 2,10092. Pada Tabel 4.10 diketahui bahwa hanya variabel K_1, K_2 yang berpengaruh signifikan terhadap variabel terikat. Sedangkan diperoleh nilai VIF sebesar 1,000 artinya tidak ada korelasi antar variabel komponen utama. Oleh karena itu, dilakukan regresi ulang dengan tidak melibatkan K_3. Diperoleh hasil

$$Y = 64{,}104 + 1{,}931\ K_1 - 0{,}902\ K_2 \tag{4.3}$$

Setelah diperoleh model pada persamaan (4.3) selanjutnya, dilakukan uji F untuk mengetahui signifikansi model regresi. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : w_0 = \cdots = w_2 = 0$ (Model regresi tidak signifikan)

$H_1 : \exists w_j \neq 0,\ j = 0{,}1{,}2$ (Model regresi signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}$. Adapun hasil uji F menggunakan persamaan (2.47) adalah 53,408, sedangkan dengan $\alpha = 5\%$ diperoleh $F_{(0{,}05;2;19)} = 3{,}52$. Karena $F_{hitung} > F_{tabel}$ maka H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa model regresi berpengaruh secara signifikan.

Uji signifikansi parameter selanjutnya yaitu uji t untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : w_j = 0,\ j = 0,1,2$ (Koefisien regresi ke- j tidak signifikan)

$H_1 : w_j \neq 0,\ j = 0, 1,2$ (Koefisien regresi ke- j signifikan)

Kriteria uji t yaitu tolak H_0 jika $|t_{hitung}| > t_{tabel}$. Nilai t_{hitung} diperoleh menggunakan persamaan (2.48). Adapun diperoleh hasil pada Tabel 4.11.

Tabel 4.11 Perulangan Uji t Metode *Principal Component Regression*

Variabel	$\boldsymbol{t_{hitung}}$	$\boldsymbol{\lvert t_{hitung}\rvert}$	**Keputusan**	**VIF**
Konstanta	173,910	173,910	Signifikan	
K_1	9,764	9,764	Signifikan	1,000
K_2	-3,390	3,390	Signifikan	1,000

Sedangkan, nilai t_{tabel} diperoleh sebesar 2,09302. Dapat disimpulkan bahwa variabel K_1 dan K_2 berpengaruh signifikan terhadap variabel terikat. Dan diperoleh nilai VIF kurang dari 10 yaitu bernilai 1,000 artinya tidak ada korelasi antar variabel komponen utama. Dengan demikian, dapat disimpulkan bahwa estimasi dengan metode *Principal Component Regression* (PCR) tidak mengandung multikolinieritas.

4.2.5 Mengembalikan Persamaan Regresi dalam Variabel Standar, Z

Setelah diperoleh model PCR terstandarisasi, maka langkah berikutnya dilakukan substitusi persamaan (4.1) ke dalam (4.3) sebagai berikut:

$$Y = 64{,}104 + 1{,}931\ K_1 - 0{,}902\ K_2$$

$$Y = 64{,}104 + 1{,}931\,(0{,}402\,Z_1 + 0{,}379\,Z_2 - 0{,}323\,Z_3 + 0{,}440\,Z_4 + 0{,}198\,Z_5 + 0{,}348\,Z_6 + 0{,}244\,Z_7 + 0{,}422\,Z_8) - 0{,}902(0{,}417\,Z_1 - 0{,}366\,Z_2 + 0{,}206\,Z_3 - 0{,}261\,Z_4 - 0{,}068\,Z_5 + 0{,}503\,Z_6 + 0{,}483\,Z_7 - 0{,}300\,Z_8)$$

$$\hat{Y} = 64{,}104 + 0{,}4001\,Z_1 + 1{,}0619\,Z_2 - 0{,}8096\,Z_3 + 1{,}0850\,Z_4 + 0{,}4437\,Z_5 + 0{,}2183\,Z_6 + 0{,}0355\,Z_7 + 1{,}0854\,Z_8 \qquad (4.4)$$

4.2.6 Mengembalikan Persamaan Regresi dalam Variabel Awal, *X*

Langkah terakhir adalah mengembalikan persamaan regresi dalam variabel awal. Diperoleh hasil sebagai berikut:

$$\hat{Y} = 64{,}104 + 0{,}4001\left(\frac{X_1 - 242{,}0709}{101{,}38459}\right) + 1{,}0619\left(\frac{X_2 - 4{,}1518}{1{,}88571}\right) - 0{,}8096\left(\frac{X_3 - 21{,}6173}{7{,}26240}\right) + 1{,}0850\left(\frac{X_4 - 11704{,}95}{6082{,}677}\right) + 0{,}4437\left(\frac{X_5 - 66{,}1545}{2{,}17721}\right) + 0{,}2183\left(\frac{X_6 - 123907{,}05}{53725{,}697}\right) + 0{,}0355\left(\frac{X_7 - 82{,}05}{35{,}435}\right) + 1{,}0854\left(\frac{X_8 - 7{,}5236}{1{,}10359}\right)$$

$$\hat{Y} = 39{,}6656 + 0{,}0039\,X_1 + 0{,}5631\,X_2 - 0{,}1115\,X_3 + 0{,}0001\,X_4 + 0{,}2038\,X_5 + 0{,}000004\,X_6 + 0{,}0010\,X_7 + 0{,}9835\,X_8 \qquad (4.5)$$

Interpretasi berdasarkan model (4.5) diketahui bahwa konstanta sebesar 39,6656 menunjukkan bahwa ketika semua variabel bernilai nol maka nilai dari IPM sebesar 39,6656. Koefisien regresi variabel jumlah penduduk (X_1) sebesar 0,0039 menunjukkan bahwa setiap kenaikan satu satuan jumlah penduduk (X_1)

maka IPM akan meningkat sebesar 0,0039 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel TPT (X_2) sebesar 0,5631 menunjukkan bahwa setiap kenaikan satu satuan TPT (X_2) maka IPM akan meningkat sebesar 0,5631 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel persentase penduduk miskin (X_3) sebesar -0,1115 menunjukkan bahwa setiap kenaikan satu satuan penduduk miskin (X_3) maka IPM akan menurun sebesar 0,1115 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel PDRB (X_4) sebesar 0,0001 menunjukkan bahwa setiap kenaikan satu satuan PDRB (X_4) maka IPM akan meningkat sebesar 0,0001 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel AHH (X_5) sebesar 0,2038 menunjukkan bahwa setiap kenaikan satu satuan AHH (X_5) maka IPM akan meningkat sebesar 0,2038 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel penduduk usia 15 tahun yang bekerja (X_6) sebesar 0,000004 menunjukkan bahwa setiap kenaikan satu satuan penduduk usia 15 tahun yang bekerja (X_6) maka IPM akan meningkat sebesar 0,000004 satuan, dengan asumsi bahwa variabel bebas lainnya tetap.Koefisien regresi variabel fasilitas kesehatan (X_7) sebesar 0,0010 menunjukkan bahwa setiap kenaikan satu satuan sarana kesehatan (X_7) maka IPM akan meningkat sebesar 0,0010 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel rata-rata lama sekolah (X_8) sebesar 0,9835 menunjukkan bahwa setiap kenaikan satu satuan rata-rata lama sekolah (X_8) maka IPM akan meningkat sebesar 0,9835 satuan, dengan asumsi bahwa variabel bebas lainnya tetap.

4.3 Regresi *Ridge*

4.3.1 Metode Pemusatan dan Penskalaan

Metode pemusatan dan penskalaan dilakukan menggunakan persamaan (2.10) dan (2.11). Dalam transformasi data ini membutuhkan nilai *mean* dan *standard deviation* dari masing-masing variabel menggunakan persamaan (2.12) dan (2.13). Adapun hasil transformasi data menggunakan metode pemusatan dan penskalaan dapat dilihat pada lampiran 2.

4.3.2 Estimasi Parameter Regresi *Ridge*

Setelah diperoleh data dalam bentuk standarisasi, maka dilakukan estimasi regresi *ridge* terlebih dahulu menentukan tetapan bias c dengan menggunakan estimator parameter *Hoerl and Kennard* pada persamaan (2.46). Sehingga diperoleh $\widehat{\boldsymbol{\alpha}}_{\mathbf{R}}$, $\widehat{\boldsymbol{\beta}}_{\mathbf{R}}^{*}$, dan VIF yang disajikan pada tabel di bawah ini.

Tabel 4. 12 Estimator Parameter *Hoerl and Kennard*

Nilai c_{HK}	$\widehat{\boldsymbol{\alpha}}_{\mathbf{R}}$	$\widehat{\boldsymbol{\beta}}_{\mathbf{R}}^{*}$	**VIF**
0,7800	-0,0028	0,001	0,1656
	-0,3357	-0,081	3,2981
	-0,3190	-0,093	1,5887
	0,2673	0,538	6,0630
	-0,1224	0,369	1,3908
	-0,1472	-0,078	0,5102
	0,2132	0,160	2,3832
	0,4564	0,321	4,2932

Pada Tabel 4.12 diperoleh nilai $c_{HK} = 0{,}7800$ dan nilai VIF kurang dari 10. Nilai VIF tersebut diperoleh menggunakan persamaan berikut.

$$VIF = \left(\frac{1}{n-1}(\boldsymbol{X}'\boldsymbol{X}) + \boldsymbol{Q}c\boldsymbol{Q}'\right)^{-1}\left(\frac{1}{n-1}(\boldsymbol{X}'\boldsymbol{X})\right)\left(\frac{1}{n-1}(\boldsymbol{X}'\boldsymbol{X}) + \boldsymbol{Q}c\boldsymbol{Q}'\right)^{-1} \quad (4.6)$$

Dengan demikian, dapat disimpulkan bahwa estimasi dengan metode *Hoerl and Kennard* tidak mengandung multikolinieritas.

4.3.3 Uji Signifikansi Parameter

Setelah diperoleh estimasi dengan metode *Hoerl and Kennard* selanjutnya, dilakukan uji F untuk mengetahui signifikansi model regresi. Adapun hipotesis yang digunakan sebagai berikut:

H_0 : $\beta^*_{R(1)}$=... $= \beta^*_{R(8)} = 0$ (Model regresi tidak signifikan)

H_1 : $\exists \beta^*_{R(j)} \neq 0,\ j = 1, \dots, 8$ (Model regresi signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}$. Adapun hasil uji F menggunakan persamaan (2.47) adalah 25,2968, sedangkan dengan $\alpha = 5\%$ diperoleh $F_{(0,05;8;13)} = 2{,}77$. Karena $F_{hitung} > F_{tabel}$ maka H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa model regresi *ridge* berpengaruh secara signifikan.

Uji signifikansi parameter berikutnya yaitu uji t untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Adapun hipotesis yang digunakan sebagai berikut:

H_0: $\beta^*_{R(j)} = 0$ untuk $j = 1, \dots, 8$ (Koefisien regresi ke-j tidak signifikan)

H_1: $\beta^*_{R(j)} \neq 0$ untuk $j = 1, \dots, 8$ (Koefisien regresi ke-j signifikan)

Kriteria uji t yaitu jika $|t_{hitung}| > t_{tabel}$ maka H_0 ditolak. Berikut nilai t_{hitung} yang diperoleh menggunakan persamaan (2.48) dengan bantuan *software* MATLAB diperoleh hasil pada Tabel 4.13.

Tabel 4.13 Uji t Regresi *Ridge*

Variabel	t_{hitung}	$\lvert t_{hitung} \rvert$	t_{tabel}	**Keputusan**
Z_1	0,002	0,002		Tidak Signifikan
Z_2	-0,658	0,658		Tidak Signifikan
Z_3	-0,942	0,942		Tidak Signifikan
Z_4	2,764	2,764	2,16037	Signifikan
Z_5	4,578	4,578		Signifikan
Z_6	-0,129	0,129		Tidak Signifikan
Z_7	1,467	1,467		Tidak Signifikan
Z_8	2,196	2,196		Signifikan

Pada Tabel 4.13 diketahui bahwa variabel Z_4 yaitu PDRB, variabel Z_5 yaitu AHH, dan variabel Z_8 yaitu rata-rata lama sekolah berpengaruh signifikan terhadap IPM. Sedangkan, variabel Z_1 yaitu jumlah penduduk, variabel Z_2 yaitu TPT, variabel Z_3 yaitu persentase penduduk miskin, variabel Z_6 yaitu penduduk usia 15 tahun yang bekerja, dan variabel Z_7 yaitu jumlah fasilitas kesehatan tidak berpengaruh signifikan terhadap IPM. Karena terdapat variabel bebas Z_1, Z_2 Z_3, Z_6, Z_7 yang tidak berpengaruh signifikan terhadap variabel terikat maka dilakukan regresi ulang dengan menghilangkan satu per satu variabel bebas hingga diperoleh variabel yang signifikan. Adapun hasilnya pada Tabel 4.14.

Tabel 4.14 Perulangan Estimator Parameter *Hoerl and Kennard*

Variabel	**Nilai** c_{hk}	$\hat{\alpha}_R$	$\hat{\beta}_R^*$	**VIF**
Z_4		0,0405	0,4684	1,8908
Z_5	0,2760	-0,1116	0,3291	1,0595
Z_8		0,6843	0,3933	3,3629

Dengan nilai $\hat{\boldsymbol{\beta}}_R^*$ di atas, Selanjutnya, dilakukan uji F kembali untuk mengetahui signifikansi model regresi. Adapun hipotesis yang digunakan sebagai berikut:

$H_0 : \beta_{R(4)}^* = \beta_{R(5)}^* = \beta_{R(8)}^* = 0$ (Model regresi tidak signifikan)

$H_1 : \exists\beta^*_{R(j)} \neq 0,\ j = 4{,}5{,}8$ (Model regresi signifikan)

dengan kriteria uji, yaitu H_0 ditolak jika $F_{hitung} > F_{tabel}$. Adapun hasil uji F menggunakan persamaan (2.47) adalah 26,1687, sedangkan dengan $\alpha = 5\%$ diperoleh $F_{(0{,}05;3;18)} = 3{,}16$. Karena $F_{hitung} > F_{tabel}$ maka H_0 ditolak. Dengan demikian, dapat disimpulkan bahwa model regresi *ridge* berpengaruh secara signifikan.

Uji signifikansi parameter selanjutnya yaitu uji t untuk mengetahui pengaruh masing-masing variabel bebas terhadap variabel terikat. Adapun hipotesis yang digunakan sebagai berikut:

$H_0: \beta^*_{R(j)} = 0$ untuk $j = 4{,}5{,}8$ (Koefisien regresi ke-j tidak signifikan)

$H_1: \beta^*_{R(j)} \neq 0$ untuk $j = 4{,}5{,}8$ (Koefisien regresi ke-j signifikan)

Kriteria uji t yaitu tolak H_0 jika $|t_{hitung}| > t_{tabel}$. Berikut nilai t_{hitung} yang diperoleh menggunakan persamaan (2.48) dengan $\alpha = 5\ \%$ yang disajikan pada Tabel 4.15.

Tabel 4.15 Perulangan Uji t Regresi *Ridge*

Variabel	$\boldsymbol{t_{hitung}}$	$\boldsymbol{\lvert t_{hitung}\rvert}$	**Keputusan**
Z_4	3,9153	3,9153	Signifikan
Z_5	4,7801	4,7801	Signifikan
Z_8	3,3408	3,3408	Signifikan

Sedangkan, nilai t_{tabel} diperoleh sebesar 2,10092. Karena $|t_{hitung}| > t_{tabel}$ maka tolak H_0 artinya variabel Produk Domestik Regional Bruto (Z_4), Angka Harapan Hidup (Z_5), dan rata-rata lama sekolah (Z_8) berpengaruh signifikan terhadap variabel IPM.

4.3.4 Transformasi Parameter Regresi *Ridge*

Nilai estimasi parameter regresi *ridge* standar yang telah diperoleh ditransformasi ke bentuk awal menggunakan persamaan (2.15) dan (2.16) sehingga diperoleh hasil sebagai berikut:

$$\hat{\beta}_{R(4)} = \left(\frac{S_y}{S_4}\right)\hat{\beta}^*_{R(4)}$$

$$= \left(\frac{4{,}23184}{6082{,}677}\right)(0{,}4684)$$

$$= 0{,}0003$$

$$\hat{\beta}_{R(5)} = \left(\frac{S_y}{S_5}\right)\hat{\beta}^*_{R(5)}$$

$$= \left(\frac{4{,}23184}{2{,}17721}\right)(0{,}3291)$$

$$= 0{,}6396$$

$$\hat{\beta}_{R(8)} = \left(\frac{S_y}{S_8}\right)\hat{\beta}^*_{R(8)}$$

$$= \left(\frac{4{,}23184}{1{,}10359}\right)(0{,}3933)$$

$$= 1{,}5081$$

$$\hat{\beta}_{R(0)} = \bar{Y} - \hat{\beta}_{R(4)}\bar{X}_4 - \hat{\beta}_{R(5)}\bar{X}_5 - \hat{\beta}_{R(8)}\bar{X}_8 = 6{,}6308$$

Sehingga diperoleh model regresi *ridge* menggunakan metode *Hoerl and Kennard* sebagai berikut.

$$\hat{Y}_R = 6{,}6308 + 0{,}0003\, X_4 + 0{,}6396\, X_5 + 1{,}5081\, X_8 \tag{4.7}$$

Interpretasi berdasarkan model (4.7) adalah konstanta sebesar 6,630 menunjukkan bahwa ketika semua variabel bernilai nol maka nilai dari IPM sebesar 6,630. Koefisien regresi variabel PDRB (X_4) sebesar 0,0003 menunjukkan bahwa setiap kenaikan satu satuan PDRB (X_4) maka IPM akan meningkat sebesar 0,0003 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel AHH (X_5) sebesar 0,6396 menunjukkan bahwa setiap kenaikan satu satuan AHH (X_5) maka IPM akan meningkat sebesar 0,6396 satuan, dengan asumsi bahwa variabel bebas lainnya tetap. Koefisien regresi variabel rata-rata lama sekolah (X_8) sebesar 1,5081 menunjukkan bahwa setiap kenaikan satu satuan rata-rata lama sekolah (X_8) maka IPM akan meningkat sebesar 1,5081 satuan, dengan asumsi bahwa variabel bebas lainnya tetap.

4.4 Kebaikan Model

Dari pembahasan di atas, nilai koefisien penduga dan $\boldsymbol{t_{hitung}}$ dari kedua metode dapat dilihat pada Tabel 4.16.

Tabel 4.16 Nilai Koefisien Penduga dan $\boldsymbol{t_{hitung}}$ dari Kedua Metode

Koefisien Penduga	**Nilai Koefisien Penduga**		**Nilai $\boldsymbol{t_{hitung}}$**	
	Principal Component Regression	**Regresi *Ridge***	***Principal Component Regression***	**Regresi *Ridge***
β_1	0,400	0,001	12,396	0,002
β_2	1,0619	-0,081	36,553	-0,658
β_3	-0,8096	-0,093	-40,687	-0,942
β_4	1,0850	0,538	41,225	2,764
β_5	0,4437	0,369	43,511	4,578
β_6	0,2183	-0,078	6,136	-0,129
β_7	0,0355	0,160	1,093	1,467
β_8	1,0854	0,321	39,757	2,196

Berdasarkan Tabel 4.16 diketahui bahwa koefisien penduga pada *principal component regression* menunjukkan semua variabel bebas berpengaruh nyata pada taraf signifikansi 0,05, kecuali variabel X_7. Hal ini ditunjukkan dengan

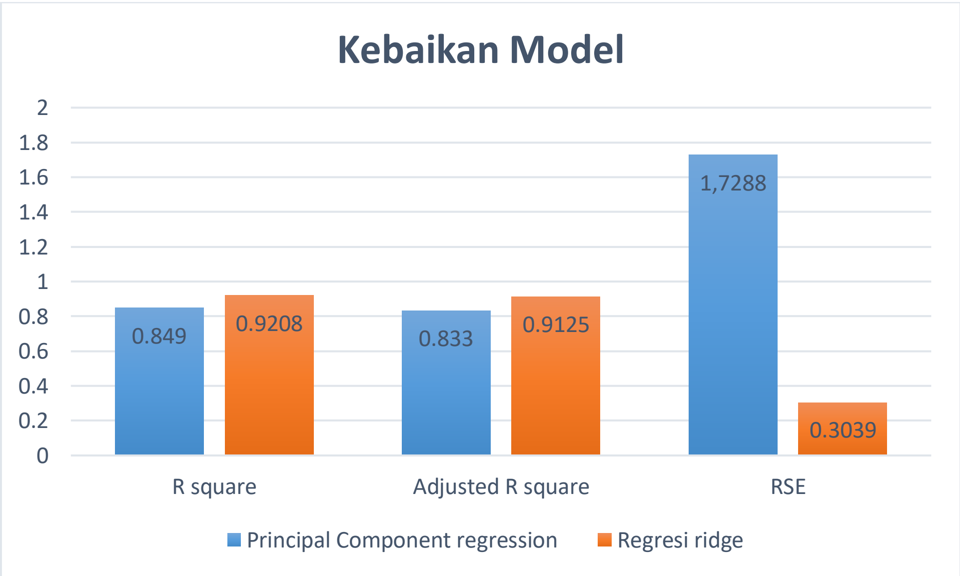
$|t_{hitung}| > t_{tabel}$ dengan t_{tabel} = 2,16037 yang artinya H_0 ditolak sehingga koefisien regresi berpengaruh signifikan. Sedangkan pada regresi *ridge* menunjukkan bahwa koefisien regresi yang berpengaruh signifikan adalah variabel X_4, X_5, X_8. Sedangkan variabel lainnya tidak berpengaruh signifikan.

Model yang telah dilakukan uji signifikansi parameter akan dilihat kebaikan model berdasarkan nilai R^2, R_{adj}, dan RSE. Nilai R^2 dihitung menggunakan persamaan (2.49), nilai R_{adj} dihitung menggunakan persamaan (2.50), dan RSE dihitung menggunakan persamaan (2.51). Adapun hasilnya dalam bentuk Tabel 4.17.

Tabel 4.17 Nilai $\boldsymbol{R^2}, \boldsymbol{R_{adj}}$, dan RSE

Metode	$\boldsymbol{R^2}$	$\boldsymbol{R_{adj}}$	**RSE**
Principal Component Regression	0,849	0,833	1,7288
Regresi *Ridge*	0,9208	0,9125	0,3039

Apabila digambarkan dalam histogram, hasilnya dapat dilihat pada Gambar 4.1.



Gambar 4.1 Histogram Kebaikan Model

Berdasarkan Gambar 4.1, dapat diketahui bahwa nilai R^2 dan R_{adj} dari regresi *ridge* dengan parameter *Hoerl and Kennard* memiliki nilai lebih tinggi daripada *principal component regression.* Sedangkan, nilai RSE regresi *ridge* dengan parameter *Hoerl and Kennard* memiliki nilai lebih kecil daripada *principal component regression.*

Berdasarkan Gambar 4.1 diketahui bahwa nilai R_{adj} dari *principal component regression* sebesar 83,3 % artinya terdapat hubungan antara variabel terikat dengan variabel bebas yang menjelaskannya, sedangkan 16,7 % dipengaruhi oleh variabel bebas lainnya. Selanjutnya RSE *principal component regression* diperoleh sebesar 1,7288. Sedangkan R_{adj} dari regresi *ridge* sebesar 91,25 % artinya terdapat hubungan antara variabel terikat dengan variabel bebas yang menjelaskannya, sedangkan 8,75 % dipengaruhi oleh variabel bebas lainnya. Kemudian RSE regresi *ridge* diperoleh sebesar 0,3039. Dari kedua metode diperoleh selisih R_{adj} sebesar 7,95 % yang lebih besar dari α dengan α sebesar 5%.

4.5 Manfaat Bekerja dalam Perspektif Al-Qur'an

Manusia dilahirkan memiliki banyak kebutuhan dan keinginan. Kebutuhan dan keinginan dapat terpenuhi apabila seseorang sungguh-sungguh bekerja. Tanpa adanya usaha maka keinginan seseorang atau kebutuhan akan sulit terpenuhi. Dalam usaha, islam memberikan keleluasaan kepada umatnya selama tidak menyimpang dari prinsip-prinsip syariat islam. Oleh karena itu, sangat masuk akal jika islam memandang bekerja sebagai kewajiban setiap orang muslim seperti dalam Q.S Al Insyirah ayat 7.

فَاِذَا فَرَغْتَ فَانْصَبْۙ ٧

Artinya:*" Apabila engkau telah selesai (dengan suatu kebajikan), teruslah bekerja keras (untuk kebajikan yang lain)."*

Menurut Muthahari, ayat di atas menjelaskan bahwa apabila seseorang telah selesai pada satu pekerjaan, maka sebaiknya kerjakan pekerjaan yang lain dan harapkanlah rahmat dari Tuhanmu. Jangan biarkan diri sendiri tenggelam dalam kenyamanan. Islam memandang bekerja bukan hanya sekadar memenuhi kebutuhan perut tetapi juga untuk memelihara harga diri. Seseorang yang yang telah bekerja dan bersungguh-sungguh dalam pekerjaannya akan bertambah martabat dan kemuliaannya. Tidak dibenarkan seorang muslim hanya berdoa saja mengharap rezeki tanpa ada keinginan untuk berusaha. Tetapi tidak dibenarkan pula untuk mengandalkan kemampuan sehingga melupakan pertolongan Allah karena akan menimbulkan rasa sombong atas dirinya sendiri.

Selain itu, manfaat dari bekerja adalah untuk meneguhkan syariat islam. Seseorang yang menggunakan kemampuannya baik yang bersifat jasmani atau rohani untuk mendapatkan karunia Allah dalam rangka memenuhi kebutuhan hidupnya dan keluarga yang menjadi tanggungannya maka ia sedang *jihad fi sabilillah*. Kemudian, motivasi kerja seorang muslim adalah meraih cinta dari Allah SWT. Untuk meraih cinta Allah SWT, maka seseorang muslim dituntut untuk menggunakan setiap kemampuan dan kesempatan yang ada untuk menghasilkan sesuatu yang bermanfaat.

Implementasi model pada penelitian ini mempunyai hasil yang signifikan pada semua variabel, kecuali variabel fasilitas kesehatan pada metode PCR. Sedangkan, metode regresi *ridge* variabel yang signifikan adalah variabel Produk

Domestik Regional Bruto (PDRB), Angka Harapan Hidup (AHH), dan rata-rata lama sekolah serta mempunyai hasil keakuratan yang baik pada kedua metode. Sehingga pemodelan ini dapat digunakan sebagai pertimbangan untuk Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur (NTT) untuk menyusun kebijakan-kebijakan yang dapat meningkatkan indeks pembangunan manusia di provinsi tersebut. Oleh karena itu, model *principal component regression* dan regresi *ridge* yang diimplementasikan pada data IPM akan bermanfaat bagi pihak BPS Provinsi Nusa Tenggara Timur.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dari pembahasan, dapat ditarik kesimpulan yaitu berdasarkan nilai koefisien penduga dan t_{hitung} dari metode *principal component regression* menunjukkan semua variabel bebas berpengaruh nyata pada taraf signifikansi 0,05, kecuali variabel X_7. Sedangkan, pada regresi *ridge* menunjukkan bahwa koefisien regresi yang berpengaruh signifikan adalah variabel X_4, X_5, X_8. Selanjutnya, diperoleh hasil keakuratan model *principal component regression* adalah R^2 sebesar 0,849 kemudian R_{adj} sebesar 0,833 dan RSE sebesar 1,7288. Sedangkan, hasil keakuratan model regresi *ridge* dengan parameter *Hoerl and Kennard* adalah R^2 sebesar 0,9208 kemudian R_{adj} sebesar 0,9125 dan RSE sebesar 0,3039. Dari kedua metode diperoleh selisih R_{adj} sebesar 7,95 % yang lebih besar dari α. Oleh karena itu, dapat disimpulkan bahwa model regresi *ridge* lebih baik daripada *principal component regression* dalam mengatasi multikolinieritas pada data IPM di Nusa Tenggara Timur. Adapun model regresi *ridge* adalah $\hat{Y}_R = 6{,}6308 + 0{,}0003\, X_4 + 0{,}6396\, X_5 + 1{,}5081\, X_8$.

5.2 Saran untuk Penelitian Lanjutan

Untuk penelitian selanjutnya diharapkan dapat memodelkan data IPM yang mengandung multikolinieritas menggunakan perbandingan metode lain dan parameter *ridge* yang lain.

DAFTAR PUSTAKA

Alydrus, M. (2009). *Agar Hidup Selalu Berkah.* Bandung: Mizan Pustaka.

Arsyad, L. (1997). *Ekonomi Pembangunan.* Yogyakarta: STIE YKPN.

Aziz, A. (2010). *Ekonometrika: Teori & Praktik Eksperimen dengan Matlab.* Malang: UIN Maliki Press.

Blanchard, Oliver and David R. Johnson. (2013). *Macroeconomics. Sixth Edition.* United States of America :Pearson.

Berg, Hendrik Van den. (2005). *Economic Growth and Development.* Singapura: McGraw-Hill.

Departemen Agama RI. (2015). *Al-Qur'an dan Terjemahannya*. Jakarta: CV Darus Sunnah.

Dewi, N.L.S., & Sutrisna, I.K. (2014). Pengaruh Komponen Indeks Pembangunan Manusia Terhadap Pertumbuhan Eknomi Provinsi Bali. *E-Jurnal Ekonomi Pembangunan Universitas Udayana*, 3(3), 106-114.

Draper, H., & Smith, H. (1992). *Analisis Regresi Terapan. Edisi 2. (Terjemahan: Bambang-Sumantri).* Jakarta: PT Gramedia Pustaka Utama.

Faizia, T., Alan, P., & Hasbi, Y. (2019). Pemodelan Indeks Pembangunan Manusia Di Jawa Tengah Dengan Regresi Komponen Utama Robust. *Gaussian*, Vol 8, No.2, 253-271.

Fitrianto, A., & Yik, L.C. (2014). Performance of Ridge Regression Estimator Methods on Small Sample Size By Varying Correlation Coefficients. *Journal of Mathematics and Statistics. 10 (1)*, 25-29.

Ghozali, I. (2016). *Aplikasi Analisis Multivariate dengan Progam IBM SPSS 23.* Semarang: Badan Penerbit Universitas Diponegoro.

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics Fifth Edition.* New York: Mc Graw-Hill Irwin.

Hasriani, I. d. (2016). Perbandingan Regresi Ridge dan Principal Component Analysis Dalam Mengatasi Masalah Multikolinieritas. *Jurnal Teknosains, Volume 10 Nomor 2*, 125-135.

Hoerl, A.E., Kennard, R.W. (1970). Ridge regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1),55-67.

James, G., Witten, D., Hastie, T., dan Tibshirani, R. (2013). *An Introduction to Statistical Learning With Applications in R.* New York: Springer.

Johnson, R.A. dan Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis. Sixth edition,* . Prentice Hall: New Jersey.

Jumiati, S. M. (2018). Penerapan Analisis Komponen Utama untuk Mereduksi Variabel dalam Pengukuran Desain Helm. *Bimaster Vol.7, No.3,* 225-230.

Kuncoro, M. (2006). *Ekonomi Pembangunan, Teori, Masalah, dan Kebijakan. Edisi Ketiga.* Yogyakarta: UPP AMP YKPN.

Kutner, M. H., & Nachtsheim, N. (2005). *Applied Linear Statistical Models Fifth Edition.* New York: Mc Graw Hill.

Mankiw, G. N. (2012). *Pengantar Ekonomi Makro.* Jakarta: Erlangga.

Marcus, G. L., Lesnussa, Y.A., & Wattimanela, H.J. (2012). Analisis Komponen Utama untuk Mengatasi Multikolinieritas dalam Analisis Regresi Linier Berganda (Studi Kasus: Cuah Hujan di Kota Ambon Tahun 2010). *Jurnal Barekang Vol. 6, No. 1,* 31-40.

Mardikyan, S., Cetin, E. (2008). Efficient Choice of Biasing Constant for Ridge Regression. Int. J. Contemp. Math. Science Vol.3 No.11,.527-536.

Montgomery, D. C., Peck, E. A., dan Vining, G .G .(2012). *Introduction to Linear Regression Analysis. Fifth Edition.* . New Jersey : John Wiley & Sons, lnc.

Morrison, D. (1978). *Multivariate Statistical Methods Series in Probability and Statistics.* Singapore: Mc Graw Hill.

Mulyadi. (2003). *Ekonomi Sumber Daya Manusia Dalam Perspektif Pembangunan.* Jakarta: Rajagrafindo Persada.

Muthahari, M. (2000). *Tafsir Surah Pilihan, Mengisi Hidup dengan Surah Penuh Berkah.* Bandung: Pustaka Hidayah.

Nanga, R. (2001). *Makroekonomi: Masalah Dana Kebijakan.* Jakarta: PT. Raja Grafindo Persada.

Nurmainah, S. (2013). Analisis Pengaruh Belanja Modal Pemerintah Daerah, Tenaga Kerja Terserap dan Indeks Pembangunan Manusia Terhadap Pertumbuhan Ekonomi dan Kemiskinan (Studi Kasus 35 Kabupaten/Kota di Provinsi Jawa Tengah). *Jurnal Bisnis dan Ekonomi*, 20 (2).

Noor, M. I. (2012). *Motivasi Islam Dan Motivasi Prososial Pada Lembaga Amil Zakat.* Semarang: Fak Ekonomi dan Bisnis.

Notiragayu dan Nisa, K. (2008). Analisis Regresi Komonen Utama Robust untuk Data Mengandung Pencilan. *Jurnal Sains MIPA,*14(1), 45-50.

Nurdin, I., Sugiman., & Sunarmi. (2018). Penerapan Kombinasi Metode Ridge Regression (RR) dan Metode Generalized Least Square (GLS) untuk Mengatasi Masalah Multikolinieritas dan Autokorelasi. *Jurnal MIPA 41 (1)*, 58-68.

Ryan, T. P.(1997). *Modern Regression Methods.* New York: John Wiley & Sons.

Shihab, M. Q. (2016). *Tafsir Al-Misbah: Pesan, Kesan dan Keserasian Al-Qur'an.* Jakarta: Lentera Hati.

Siagaan, D. Sugiarto. (2006). *Metode Statistika Untuk Bisnis dan Ekonomi.* Jakarta: PT Gramedia Pustaka Utama.

Solekakh, N.A., Ispriyanti, D., & Sudarno. (2015). Estimasi Parameter Regresi Ridge Menggunakan Iterasi Hoerl, Kennard, dan Baldwin (HKB) untuk penanganan Multikolinieritas. *Jurnal Gaussian, Volume 4 No.4*, 1109-1116.

Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan Multikolinieritas Dengan Menggunakan Analisis Regresi Komponen Utama Pada Kasus Impor Beras di Provinsi Sulut. *Jurnal Ilmiah Sains,* Vol.18 No.1, 18-24.

Sukirno, Sadono. (1985). *Ekonomi Pembangunan: Proses, Masalah, dan Kebijaksanaan.* Jakarta: LPFE-UI.

Sukmono, A. (2014). *Penggunaan Partial Least Square Regression (PLSR) untuk Mengatasi Multikolinieritas Dalam Estimasi Klorofil Daun Tanaman Padi dengan Citra Hiperspektral.* Program Studi Teknik Geodesi, Fakultas Teknik, Universitas Diponegoro.

Sulistyono & Sulistiyowati, W. (2017). "Peramalan Produksi dengan Metode Regresi Linier Berganda". *Prozima*, Vol. 1 (2), 82-89.

Suliswanto, M.S.W. (2010). Pengaruh Produk Domestik Bruto (PDB) dan Indeks Pembangunan Manusia (IPM) Terhadap Angka Kemiskinan di Indonesia. *Jurnal Ekonomi Pembangunan*, 8(2), 357-366.

Suma, M. A. (2015). *Tafsir Ayat Ekonomi (Teks, Terjemah, dan Tafsir).* Jakarta: Amzah.

Supriatna, N., Ruhimat, M.& Kosim. (2006). *Ilmu Pengetahuan Sosial (Geografi, Sejarah, Sosiologi, Ekonomi).* Bandung:Grafindo Media Pratama.

Supriyadi, E. M. (2017). Perbandingan Metode Partial Least Square (PLS) dan Principal Component Regression (PCR) untuk Mengatasi Multikolinieritas pada Model Regresi Linear Berganda. *Unnes Journal of Mathematics*, 6(2), 117-128.

Suryanto. (1988). *Metode Statistik Multivariat I.* P2LPTK: Jakarta.

Tirink, C., Abaci, S. H., & Onder, H. (2020). Comparison Of Ridge Regression and Least Square Methods In The Presence of Multicollinearity for Body Measurements in Saanen Kids. *Journal of the Institute of Science and Technology*, 10(2), 1429-1437.

Tjiptoherijanto, P. (1996). *Sumber Daya Manusia dalam Pembangunan Nasional.* Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.

UI-Sulfie, A., Yahya, A. S., & Ramli, N. A. (2011). Improving Multiple Linear Regression Model Using Principal Component Analysis for Predicting PM10 Concentration in Seberang Prai, Pulau Minang. International *Journal of Environmental Sciences* Vol.2, No.2.

Youngker, J. (2012). *Ridge Estimation and its Modifications for Linear Regression with Deterministic or Stochastic Predictors.* Canada: Ottawa University.

LAMPIRAN

Lampiran 1. Data Penelitian

Kabupaten	X1	X2	X3	X4	X5
Sumba Barat	145.1	3.96	28.17	10922	67.08
Sumba Timur	244.82	3.49	29.65	15174	65.13
Kupang	366.38	4.9	22.77	11366	64.63
Timor Tengah S	455.41	2.63	27.49	10154	66.42
Timor Tengah U	259.83	4.26	22.28	11252	66.96
Belu	217.97	7.42	15.37	13301	64.61
Alor	211.87	3.11	21.09	9651	61.48
Lembata	135.93	4.88	26.14	7980	67.07
Flores Timur	276.9	3.16	10.84	13671	65.2
Sikka	321.95	4	13.12	10273	67.24
Ende	270.76	2.95	23.76	14845	65.29
Ngada	165.25	4.69	12.51	13978	68.04
Manggarai	312.86	4.09	20.34	8733	67.03
Rote Ndao	143.76	4.9	27.54	10813	64.6
Manggarai B	256.32	3.72	17.71	8087	67.38
Sumba Tengah	85.48	4.02	34.49	10249	68.38
Sumba Barat D	303.65	2.36	28	6575	68.53
Nagekeo	159.73	3.09	12.61	9367	67.13
Manggarai Timur	275.6	2.1	26.52	7044	68.04
Sabu Raijua	89.33	3.08	30.18	7784	60.64
Malaka	183.9	3.63	16.04	9644	64.97
Kota Kupang	442.76	10.9	8.96	36646	69.55

\

Kabupaten	X6	X7	X8	Y
Sumba Barat	60446	32	6.6	63.53
Sumba Timur	128933	119	7.12	65.52
Kupang	190215	195	7.38	64.32
Timor Tengah S	240284	94	6.73	62.15
Timor Tengah U	135936	76	7.81	63.53
Belu	101489	47	7.35	62.68
Alor	107236	88	8.41	61.33
Lembata	68949	54	8.22	64.74
Flores Timur	129807	83	7.71	64.22
Sikka	167687	102	6.94	65.11
Ende	142737	86	7.81	67.04
Ngada	79931	60	8.52	67.88
Manggarai	171756	106	7.37	64.54
Rote Ndao	89749	100	7.59	62.39
Manggarai B	142353	74	7.3	63.89
Sumba Tengah	33175	27	6.25	61.53
Sumba Barat D	170204	75	6.34	62.28
Nagekeo	71691	61	7.89	65.81
Manggarai Timur	157523	105	7.08	60.85
Sabu Raijua	49942	66	6.65	57.02
Malaka	87226	52	6.87	60.21
Kota Kupang	198686	103	11.58	79.71

Lampiran 2. Data Standarisasi IPM Provinsi NTT Tahun 2021

Z1	Z2	Z3	Z4	Z5
-0.95647	-0.10172	0.90228	-0.12872	0.42506
0.02712	-0.35096	1.10607	0.57032	-0.47058
1.22611	0.39676	0.15873	-0.05572	-0.70023
2.10426	-0.80703	0.80865	-0.25498	0.12192
0.17517	0.05737	0.09125	-0.07447	0.36995
-0.23772	1.73313	-0.86022	0.26239	-0.70941
-0.29788	-0.55248	-0.0726	-0.33767	-2.14703
-1.04691	0.38616	0.62276	-0.61239	0.42047
0.34353	-0.52596	-1.48398	0.32322	-0.43843
0.78788	-0.08051	-1.17004	-0.23542	0.49855
0.28297	-0.63733	0.29504	0.51623	-0.39709
-0.75772	0.2854	-1.25403	0.37369	0.86599
0.69822	-0.03278	-0.17587	-0.48859	0.4021
-0.96968	0.39676	0.81553	-0.14664	-0.71401
0.14054	-0.22899	-0.53801	-0.5948	0.56285
-1.54452	-0.0699	1.77252	-0.23936	1.02216
0.60738	-0.95021	0.87887	-0.84337	1.09105
-0.81216	-0.56309	-1.24026	-0.38436	0.44803
0.33071	-1.08809	0.67508	-0.76627	0.86599
-1.50655	-0.56839	1.17905	-0.64461	-2.53284
-0.57376	-0.27672	-0.76797	-0.33882	-0.54406
1.97948	3.57859	-1.74285	4.10034	1.55954

Z6	Z7	Z8	ZY
-1.1812	-1.4123	-0.8369	-0.1356
0.09355	1.04287	-0.3658	0.33469
1.23419	3.18762	-0.1302	0.05113
2.16613	0.33736	-0.7191	-0.4617
0.2239	-0.1706	0.25948	-0.1356
-0.4173	-0.989	-0.1573	-0.3364
-0.3103	0.16804	0.80316	-0.6554
-1.0229	-0.7915	0.631	0.15038
0.10982	0.02694	0.16887	0.0275
0.81488	0.56313	-0.5289	0.23781
0.35048	0.1116	0.25948	0.69387
-0.8185	-0.6221	0.90284	0.89237
0.89062	0.67601	-0.1392	0.10311
-0.6358	0.50668	0.06013	-0.4049
0.34334	-0.2271	-0.2026	-0.0505
-1.6888	-1.5534	-1.1541	-0.6082
0.86173	-0.1988	-1.0725	-0.4309
-0.9719	-0.5939	0.33197	0.40322
0.6257	0.64779	-0.402	-0.7689
-1.3767	-0.4528	-0.7916	-1.6739
-0.6828	-0.8479	-0.5923	-0.9201
1.39187	0.59135	3.67561	3.68784

Lampiran 3. Data Variabel Baru Komponen Utama

K1	K2	K3
-1.79496	-1.19588	0.86311
-0.1891	0.91164	-0.69027
1.58033	2.66083	-1.52334
0.72254	2.86689	0.79219
0.24875	0.01749	0.39506
0.36148	-1.57164	-0.54243
-0.60659	-0.01931	-2.08592
-0.94401	-1.40465	0.33614
0.58945	-0.0075	-0.41155
0.85658	0.98524	0.49899
0.18387	0.45654	-0.32694
0.48878	-1.81913	0.50721
0.6054	1.18371	0.36016
-0.78046	-0.38773	-1.02093
-0.02823	0.27255	0.73218
-2.57552	-1.51315	1.40893
-0.75626	1.58831	1.63366
-0.56218	-1.19414	0.37074
-0.45718	1.56525	0.92246
-2.9087	-0.51027	-2.08539
-1.03787	-0.74614	-0.17463
7.00387	-2.13891	0.04056

Lampiran 4. Output Statistika Deskriptif

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
X1	22	85.48	455.41	242.0709	101.38459
X2	22	2.10	10.90	4.1518	1.88571
X3	22	8.96	34.49	21.6173	7.26240
X4	22	6575	36646	11704.95	6082.677
X5	22	60.64	69.55	66.1545	2.17721
X6	22	33175	240284	123907.05	53725.697
X7	22	27	195	82.05	35.435
X8	22	6.25	11.58	7.5236	1.10359
Y	22	57.02	79.71	64.1036	4.23184
Valid N (listwise)	22				

Lampiran 5. Output Uji F Regresi Linier Berganda SPSS

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	19.732	8	2.467	25.297	.000[b]
	Residual	1.268	13	.098		
	Total	21.000	21			

a. Dependent Variable: Zscore(Y)

b. Predictors: (Constant), Zscore(X8), Zscore(X7), Zscore(X5), Zscore(X3), Zscore(X2), Zscore(X1), Zscore(X4), Zscore(X6)

Lampiran 6. Output Estimasi Metode OLS

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	-6.767E-15	.067	
	Zscore(X1)	.014	.628	.014
	Zscore(X2)	-.082	.124	-.082
	Zscore(X3)	-.092	.099	-.092
	Zscore(X4)	.536	.195	.536
	Zscore(X5)	.369	.081	.369
	Zscore(X6)	-.091	.610	-.091
	Zscore(X7)	.161	.109	.161
	Zscore(X8)	.322	.146	.322

a. Dependent Variable: Zscore(Y)

Lampiran 7. Output Uji t Regresi Linier Berganda & Uji Multikolinieritas SPSS

Coefficients[a]

Model		t	Sig.	Collinearity Statistics	
				Tolerance	VIF
1	(Constant)	.000	1.000		
	Zscore(X1)	.022	.983	.012	85.076
	Zscore(X2)	-.663	.519	.300	3.331
	Zscore(X3)	-.932	.368	.475	2.106
	Zscore(X4)	2.756	.016	.123	8.158
	Zscore(X5)	4.580	.001	.714	1.401
	Zscore(X6)	-.149	.884	.012	80.180
	Zscore(X7)	1.472	.165	.390	2.565
	Zscore(X8)	2.201	.046	.217	4.615

a. Dependent Variable: Zscore(Y)

Lampiran 8. Output SPSS Uji Normalitas

One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		22
Normal Parameters[a,b]	Mean	.0000000
	Std. Deviation	.24568061
Most Extreme Differences	Absolute	.104
	Positive	.101
	Negative	-.104
Test Statistic		.104
Asymp. Sig. (2-tailed)		.200[c,d]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

d. This is a lower bound of the true significance.

Lampiran 9. Output Uji Glejser SPSS

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.188	.035		5.397	.000
	Zscore(X1)	-.173	.329	-1.138	-.526	.608
	Zscore(X2)	-.116	.065	-.760	-1.777	.099
	Zscore(X3)	-.003	.052	-.022	-.065	.949
	Zscore(X4)	.099	.102	.651	.973	.349
	Zscore(X5)	.052	.042	.344	1.242	.236
	Zscore(X6)	.122	.320	.803	.383	.708
	Zscore(X7)	.006	.057	.041	.108	.916
	Zscore(X8)	.004	.077	.029	.058	.955

a. Dependent Variable: Abs_RES2

Lampiran 10. Output Uji Durbin Watson SPSS

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.969[a]	.940	.902	.31225461	1.885

a. Predictors: (Constant), Zscore(X8), Zscore(X7), Zscore(X5), Zscore(X3), Zscore(X2), Zscore(X1), Zscore(X4), Zscore(X6)

b. Dependent Variable: Zscore(Y)

Lampiran 11. Output Uji *Run Test* SPSS

Runs Test

	Unstandardized Residual
Test Value[a]	.01774
Cases < Test Value	11
Cases >= Test Value	11
Total Cases	22
Number of Runs	12
Z	.000
Asymp. Sig. (2-tailed)	1.000

a. Median

Lampiran 12. Output Minitab *EigenValue* dari Matriks Korelasi

Eigenanalysis of the Correlation Matrix

Eigenvalue	3.6372	2.0095	1.0285	0.6706	0.3003	0.2293	0.1187	0.0059
Proportion	0.455	0.251	0.129	0.084	0.038	0.029	0.015	0.001
Cumulative	0.455	0.706	0.834	0.918	0.956	0.984	0.999	1.000

Lampiran 13. Output Minitab *EigenVector*

Eigenvectors

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Z1	0.402	0.417	0.124	-0.045	-0.378	-0.023	0.043	-0.709
Z2	0.379	-0.366	-0.105	0.318	-0.090	0.709	0.316	0.005
Z3	-0.323	0.206	0.046	0.880	-0.075	-0.196	0.175	-0.052
Z4	0.440	-0.261	-0.092	0.309	-0.214	-0.230	-0.719	0.138
Z5	0.198	-0.068	0.876	0.096	0.419	0.026	-0.059	-0.008
Z6	0.348	0.503	0.126	-0.034	-0.290	-0.036	0.228	0.687
Z7	0.244	0.483	-0.380	0.130	0.690	0.178	-0.191	-0.038
Z8	0.422	-0.300	-0.188	0.016	0.249	-0.610	0.509	-0.053

Lampiran 14. Output Uji F *Principal Component Regression* SPSS

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	327.681	3	109.227	40.624	.000[b]
	Residual	48.397	18	2.689		
	Total	376.079	21			

a. Dependent Variable: Y

b. Predictors: (Constant), K3, K2, K1

Lampiran 15. Output *Principal Component Regression* SPSS

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta
1	(Constant)	64.104	.350	
	K1	1.931	.188	.870
	K2	-.902	.252	-.302
	K3	.623	.353	.149

a. Dependent Variable: Y

Lampiran 16. Output Uji t *Principal Component Regression* SPSS

Coefficients[a]

Model		T	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	183.367	.000		
	K1	10.294	.000	1.000	1.000
	K2	-3.574	.002	1.000	1.000
	K3	1.767	.094	1.000	1.000

a. Dependent Variable: Y

Lampiran 17. Output Uji F Perulangan *Principal Component Regression* SPSS

ANOVA[a]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	319.286	2	159.643	53.408	.000[b]
	Residual	56.793	19	2.989		
	Total	376.079	21			

a. Dependent Variable: Y

b. Predictors: (Constant), K2, K1

Lampiran 18. Output Uji t Perulangan *Principal Component Regression* SPSS

Coefficients[a]

Model		T	Sig.	Collinearity Statistics	
				Tolerance	VIF
1	(Constant)	173.910	.000		
	K1	9.764	.000	1.000	1.000
	K2	-3.390	.003	1.000	1.000

a. Dependent Variable: Y

Lampiran 19. Output Perulangan *Principal Component Regression* SPSS

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	64.104	.369	
	K1	1.931	.198	.870
	K2	-.902	.266	-.302

a. Dependent Variable: Y

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.921[a]	.849	.833	1.72890	1.133

a. Predictors: (Constant), K2, K1

b. Dependent Variable: Y

Lampiran 20. Script Matlab Uji Regresi *Ridge*

```
clear all;
clc;

format short

Y = xlsread('DATA15.xlsx','I2:I23');
X1 = xlsread('DATA15.xlsx','A2:A23');
X2 = xlsread('DATA15.xlsx','B2:B23');
X3 = xlsread('DATA15.xlsx','C2:C23');
X4 = xlsread('DATA15.xlsx','D2:D23');
```

```
X5 = xlsread('DATA15.xlsx','E2:E23');
X6 = xlsread('DATA15.xlsx','F2:F23');
X7 = xlsread('DATA15.xlsx','G2:G23');
X8 = xlsread('DATA15.xlsx','H2:H23');

X = [X1 X2 X3 X4 X5 X6 X7 X8];

n = length(Y);
k = 8; %banyaknya variabel bebas

Xrata = mean(X);

V1 = X(:,1)-Xrata(1);
V2 = X(:,2)-Xrata(2);
V3 = X(:,3)-Xrata(3);
V4 = X(:,4)-Xrata(4);
V5 = X(:,5)-Xrata(5);
V6 = X(:,6)-Xrata(6);
V7 = X(:,7)-Xrata(7);
V8 = X(:,8)-Xrata(8);

V = [V1 V2 V3 V4 V5 V6 V7 V8];
V

Std1= (1/(n-1))*sum(V1.^2);
Std2 = (1/(n-1))*sum(V2.^2);
Std3 = (1/(n-1))*sum(V3.^2);
Std4 = (1/(n-1))*sum(V4.^2);
Std5 = (1/(n-1))*sum(V5.^2);
Std6 = (1/(n-1))*sum(V6.^2);
Std7 = (1/(n-1))*sum(V7.^2);
Std8 = (1/(n-1))*sum(V8.^2);


Std = [Std1 Std2 Std3 Std4 Std5 Std6 Std7 Std8]
Std
%std=variansi
%sqrt(std)=standar deviasi

Xt1 = V(:,1)./sqrt(Std1);
Xt2 = V(:,2)./sqrt(Std2);
Xt3 = V(:,3)./sqrt(Std3);
Xt4 = V(:,4)./sqrt(Std4);
Xt5 = V(:,5)./sqrt(Std5);
Xt6 = V(:,6)./sqrt(Std6);
Xt7 = V(:,7)./sqrt(Std7);
Xt8 = V(:,8)./sqrt(Std8);

Xt = [Xt1 Xt2 Xt3 Xt4 Xt5 Xt6 Xt7 Xt8];
Xt
C = Xt'*Xt;

Yrata = mean(Y);
Uy = Y - Yrata;
StdY = (1/(n-1))*sum(Uy.^2);
Yt = Uy./sqrt(StdY);
Yt
```

```
[Q E] = eig(C);
lambda = Q'*C*Q;
W = Xt*Q;
lambda1 = W'*W;
alphaols = ((W'*W)^-1)*W'*Yt
betaols = ((Xt'*Xt)^-1)*Xt'*Yt
VIFols = inv((1/(n-1))*C)

mse = ((Yt'*Yt)-(alphaols'*W'*Yt))/(n-k-1)
sigmakuadratols = mse;

i = 0
c1 = sigmakuadratols/alphaols(1)'*alphaols(1);
c2 = sigmakuadratols/alphaols(2)'*alphaols(2);
c3 = sigmakuadratols/alphaols(3)'*alphaols(3);
c4 = sigmakuadratols/alphaols(4)'*alphaols(4);
c5 = sigmakuadratols/alphaols(5)'*alphaols(5);
c6 = sigmakuadratols/alphaols(6)'*alphaols(6);
c7 = sigmakuadratols/alphaols(7)'*alphaols(7);
c8 = sigmakuadratols/alphaols(8)'*alphaols(8);

ci = [c1 c2 c3 c4 c5 c6 c7 c8];
K = k*ci'
I = eye(1,8);
alpharidge = ((W'*W + K*I)^-1)*W'*Yt
betaridge = Q*alpharidge

SSR = alpharidge'*(W'*Yt)
SST = (Yt'*Yt)
SSE = SST-SSR
MSR = SSR/k

alpharidgea = alphaols;
alpharidgeb = alpharidge;
mseridge = SSE/(n-k-1)
Fhitridge = (SSR/k)/(mseridge)
VIFridge = (inv((1/(n-1))*C + Q*K*I*Q'))*((1/(n-
1)).*C)*(inv((1/(n-1))*C + Q*K*I*Q'))
err = abs((alpharidgeb'*alpharidgeb)-(alpharidgea'*alpharidgea)
while(err >= 0.001)
 i = i + 1
 alpharidgea = alpharidgeb;
 c1 = (sigmakuadratols)/alpharidgeb(1)'*alpharidgeb(1);
 c2 = (sigmakuadratols)/alpharidgeb(2)'*alpharidgeb(2);
 c3 = (sigmakuadratols)/alpharidgeb(3)'*alpharidgeb(3);
 c4 = (sigmakuadratols)/alpharidgeb(4)'*alpharidgeb(4);
 c5 = (sigmakuadratols)/alpharidgeb(5)'*alpharidgeb(5);
 c6 = (sigmakuadratols)/alpharidgeb(6)'*alpharidgeb(6);
 c7 = (sigmakuadratols)/alpharidgeb(7)'*alpharidgeb(7);
 c8 = (sigmakuadratols)/alpharidgeb(8)'*alpharidgeb(8);

 ci = [c1 c2 c3 c4 c5 c6 c7 c8];
 K = k*ci'
 alpharidgeb = ((W'*W + K*I)^-1)*W'*Yt
 SSR = alpharidgeb'*(W'*Yt)
 SST = (Yt'*Yt)
 SSE = SST-SSR
 MSR = SSR/k
```

```
 mseridgeb = SSE/(n-k-1)
 koefdetridge = SSR/SST
 adjustedr2=1-((SSE/(n-k))/(SST/(n-1)))
 Fhitridge2 = (SSR/k)/mseridgeb
 err = abs((alpharidgeb'*alpharidgeb)-(alpharidgea'*alpharidgea))
 VIFridge = (inv((1/(n-1))*C + Q*K*I*Q'))*((1/(n-
1)).*C)*(inv((1/(n-1))*C + Q*K*I*Q'))
end
A = W'*W+K*I;
alpharidge = alpharidgeb
betaridge = Q*alpharidge
biasbetaridge = -Q*(inv(A))*K*I*Q'*betaols

D=inv(Xt'*Xt)

varridge=SSE/(n-(k+1))
stdridge=sqrt(varridge)
sebridge1=stdridge*sqrt(D(1,1))
sebridge2=stdridge*sqrt(D(2,2))
sebridge3=stdridge*sqrt(D(3,3))
sebridge4=stdridge*sqrt(D(4,4))
sebridge5=stdridge*sqrt(D(5,5))
sebridge6=stdridge*sqrt(D(6,6))
sebridge7=stdridge*sqrt(D(7,7))
sebridge8=stdridge*sqrt(D(8,8))

sebridge=[sebridge1; sebridge2; sebridge3; sebridge4; sebridge5;
sebridge6; sebridge7; sebridge8;]

%thit
thit=[betaridge(1)/sebridge1;
betaridge(2)/sebridge2;betaridge(3)/sebridge3;
betaridge(4)/sebridge4; betaridge(5)/sebridge5;
betaridge(6)/sebridge6;
betaridge(7)/sebridge7; betaridge(8)/sebridge8; ]

%retransformasi
beta1=(sqrt(StdY)/sqrt(Std1))*betaridge(1)
beta2=(sqrt(StdY)/sqrt(Std2))*betaridge(2)
beta3=(sqrt(StdY)/sqrt(Std3))*betaridge(3)
beta4=(sqrt(StdY)/sqrt(Std4))*betaridge(4)
beta5=(sqrt(StdY)/sqrt(Std5))*betaridge(5)
beta6=(sqrt(StdY)/sqrt(Std6))*betaridge(6)
beta7=(sqrt(StdY)/sqrt(Std7))*betaridge(7)
beta8=(sqrt(StdY)/sqrt(Std8))*betaridge(8)


beta0=Yrata-beta1*Xrata(1)-beta2*Xrata(2)-beta3*Xrata(3)-
beta4*Xrata(4)-beta5*Xrata(5)-beta6*Xrata(6)-beta7*Xrata(7)-
beta8*Xrata(8)

beta=[beta0; beta1; beta2; beta3; beta4; beta5; beta6; beta7;
beta8;]

%regresi ulang
Xt22 = [ Xt4 Xt5 Xt8];
C2 = Xt22'*Xt22;
```

```
k2=3;
I2=eye(1,3);

[Q2 E] = eig(C2);
lambda2 = Q2'*C2*Q2;
W2 = Xt22*Q2;
lambda12 = W2'*W2;
alphaols2 = ((W2'*W2)^-1)*W2'*Yt
betaols2 = ((Xt22'*Xt22)^-1)*Xt22'*Yt
VIFols2 = inv((1/(n-1))*C2)

mse2 = ((Yt'*Yt)-(alphaols2'*W2'*Yt))/(n-k2-1)
sigmakuadratols2 = mse2;

i = 0;
c42 = sigmakuadratols2/alphaols2(1)'*alphaols2(1);
c52 = sigmakuadratols2/alphaols2(2)'*alphaols2(2);
c82 = sigmakuadratols2/alphaols2(3)'*alphaols2(3);

ci2 = [ c42 c52 c82 ];
K2 = k2*ci2'
alpharidge2 = ((W2'*W2 + K2*I2)^-1)*W2'*Yt
betaridge2 = Q2*alpharidge2

SSR2 = alpharidge2'*(W2'*Yt)
SST2 = (Yt'*Yt)
SSE2 = SST2-SSR2
MSR2 = SSR2/k2

alpharidgea2 = alphaols2;
alpharidgeb2 = alpharidge2;
mseridge2 = SSE2/(n-k2-1)
Fhitridge2 = (SSR2/k)/(mseridge2)
VIFridge2 = inv((1/(n-1))*C2 + Q2*K2*I2*Q2')*(1/(n-
1)).*C2*inv((1/(n-1))*C2 +Q2*K2*I2*Q2')
err2 = abs((alpharidgeb2'*alpharidgeb2)-
(alpharidgea2'*alpharidgea2))
while(err2 >= 0.001)
 i = i + 1
 alpharidgea2 = alpharidgeb2;
 c42 = (sigmakuadratols2)/alpharidgeb2(1)'*alpharidgeb2(1);
 c52 = (sigmakuadratols2)/alpharidgeb2(2)'*alpharidgeb2(2);
 c82 = (sigmakuadratols2)/alpharidgeb2(3)'*alpharidgeb2(3);

 ci2 = [c42 c52 c82];
 K2 = k2*ci2'
 alpharidgeb2 = ((W2'*W2 + K2*I2)^-1)*W2'*Yt
 SSR2 = alpharidgeb2'*(W2'*Yt)
 SST2 = (Yt'*Yt)
 SSE2 = SST2 - SSR2
 MSR2 = SSR2/k2
 mseridge2 = SSE2/(n-k2-1)
 koefdetridge2 = SSR2/SST2
 adjustedr22=1-((SSE2/(n-k2))/(SST2/(n-1)))
 Fhitridge22 = (SSR2/k2)/mseridge2
 err2 = abs((alpharidgeb2'*alpharidgeb2)-
(alpharidgea2'*alpharidgea2))
```

```
 VIFridge2 = inv((1/(n-1))*C2 + Q2*K2*I2*Q2')*(1/(n-
1)).*C2*inv((1/(n-1))*C2 + Q2*K2*I2*Q2')
end

koefdetridge2 = SSR2/SST2
adjustedr22=1-((SSE2/(n-k2))/(SST2/(n-1)))

A2 = W2'*W2+K2*I2;
alpharidge2 = alpharidgeb2
betaridge2 = Q2*alpharidge2
biasbetaridge2 = -Q2*(inv(A2))*K2*I2*Q2'*betaols2

D2=inv(Xt22'*Xt22)

varridge2=SSE2/(n-(k2+1))
stdridge2=sqrt(varridge2)
sebridge12=stdridge2*sqrt(D2(1,1))
sebridge22=stdridge2*sqrt(D2(2,2))
sebridge32=stdridge2*sqrt(D2(3,3))

SEbridge2=[sebridge12 sebridge22 sebridge32  ]

%thit
thit=[betaridge2(1)/sebridge12; betaridge2(2)/sebridge22;
betaridge2(3)/sebridge32; ]

%retransformasi
beta12=(sqrt(StdY)/sqrt(Std4))*betaridge2(1)
beta22=(sqrt(StdY)/sqrt(Std5))*betaridge2(2)
beta32=(sqrt(StdY)/sqrt(Std8))*betaridge2(3)

beta02=Yrata-beta12*Xrata(4)-beta22*Xrata(5)-beta32*Xrata(8)

beta=[beta02; beta12; beta22; beta32;]
```

RIWAYAT HIDUP

Rizka Maulida, lahir di Gresik pada tanggal 21 Juni 2000, bertempat tinggal di Kesamben Wetan, Kecamatan Driyorejo, Kabupaten Gresik, Jawa Timur. Anak pertama dari 2 bersaudara dari pasangan Bapak Muhammad Uman dan Ibu Juwariyah.

Perempuan yang akrab disapa Rizka ini telah menempu pendidikan formal mulai dari TK Miftahul Ulum, Kemudian melanjutkan pendidikan dasar di SDN Kesamben Wetan dan lulus pada tahun 2012. Setelah itu, melanjutkan ke SMPN 1 Driyorejo dan lulus pada tahun 2015. Lalu, melanjutkan ke SMAN 1 Driyorejo dan lulus pada tahun 2018. Pada tahun yang sama melanjutkan pendidikan di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengambil program studi Matematika.

Selama menjadi mahasiswa telah mengikuti komunitas di bawah program studi matematika yaitu Mathematics English Club (MEC), Mathematics Arabic Club ((MAC), dan Serambi Aktif Matematika (SEMATA). Penulis juga pernah menjadi asisten praktikum analisis numerik 1.

KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

BUKTI KONSULTASI SKRIPSI

Nama : Rizka Maulida
NIM : 18610020
Fakultas / Jurusan : Sains dan Teknologi / Matematika
Judul Skripsi : Perbandingan *Principal Component Regression* dan Regresi *Ridge* Pada Analisis Faktor-Faktor Indeks Pembangunan Manusia
Pembimbing I : Fachrur Rozi, M.Si
Pembimbing II : Mohammad Nafie Jauhari, M.Si

No	Tanggal	Hal	Tanda Tangan
1.	24 Januari 2022	Konsultasi Bab I	1.
2.	16 Februari 2022	Konsultasi Integrasi Agama Bab I dan Bab II	2.
3.	22 Februari 2022	Revisi Integrasi Agama Bab I dan Bab II	3.
4.	25 Februari 2022	Revisi Bab I dan Konsultasi Bab III	4.
5.	10 Maret 2022	Revisi Bab III dan Konsultasi Bab II	5.
6.	16 Maret 2022	Acc Integrasi Agama Bab I dan Bab II	6.
7.	17 Maret 2022	Konsultasi Bab II	7.
8.	29 Maret 2022	Konsultasi Bab II	8.
9.	7 April 2022	Acc Bab I,II, dan III	9.
10.	12 Mei 2022	Konsultasi Bab IV	10.
11.	25 Mei 2022	Konsultasi Bab IV dan V	11.
12.	30 Mei 2022	Konsultasi Integrasi Agama Bab IV	12.
13.	31 Mei 2022	Acc Bab IV dan V	13.
14.	31 Mei 2022	Acc Integrasi Agama Bab IV	14.
15.	27 Juni 2022	Acc Keseluruhan	15.

Malang, 27 Juni 2022
Mengetahui,
Ketua Program Studi Matematika

Dr. Elly Susanti, M.Sc
NIP.197411292000122005