

**SIMULASI PENANGANAN MASALAH MULTIKOLINEARITAS
DENGAN MENGGUNAKAN METODE *PRINCIPAL COMPONENT
REGRESSION (PCR)* DAN METODE REGRESI *RIDGE***

SKRIPSI



Diajukan Kepada Fakultas Sains dan Teknologi
Universitas Islam Negeri Sunan Kalijaga Yogyakarta
Untuk Memenuhi Sebagian Syarat Memperoleh
Derajat Sarjana S-1

Disusun Oleh :

HANNIFA IRMAJIHAN NABILLA

NIM 17106010042

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2021



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi / Tugas Akhir

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Hannifa Irmajihan Nabilla
NIM : 17106010042
Judul Skripsi : Simulasi Penanganan Masalah Multikolinearitas dengan Menggunakan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge*


sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 04 Agustus 2021

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA


Mohammad Farhan Quadratullah, S.Si.,
M.Si.
NIP. 19790922 200801 1 011



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1490/Un.02/DST/PP.00.9/08/2021

Tugas Akhir dengan judul : Simulasi Penanganan Masalah Multikolinearitas dengan Menggunakan Metode Principal Component Regression (PCR) dan Metode Regresi Ridge

yang dipersiapkan dan disusun oleh:

Nama : HANNIFA IRMAJIHAN NABILLA
Nomor Induk Mahasiswa : 17106010042
Telah diujikan pada : Rabu, 11 Agustus 2021
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Mohammad Farhan Qudratullah, S.Si., M.Si
SIGNED

Valid ID: 611caff18045



Penguji I

Dr. Epha Diana Supandi, S.Si., M.Sc.
SIGNED

Valid ID: 611bdfac101b



Penguji II

Dr. Muhammad Wakhid Musthofa, S.Si.,
M.Si.
SIGNED

Valid ID: 611c35d1df55b



Yogyakarta, 11 Agustus 2021
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 611edee3587de

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Hannifa Irmajihan Nabilla
NIM : 17106010042
Program Studi : Matematika
Fakultas : Sains dan Teknologi

Dengan ini menyatakan bahwa isi skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu Perguruan Tinggi dan sesungguhnya skripsi ini merupakan hasil pekerjaan penulis sendiri sepanjang pengetahuan penulis, bukan duplikasi atau saduran dari karya orang lain kecuali bagian tertentu yang penulis ambil sebagai bahan acuan. Apabila terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab penulis.

Yogyakarta, 01 Agustus 2021



Hannifa Irmajihan Nabilla

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Dengan mengucapkan syukur atas berkat dan rahmat Allah SWT, skripsi ini saya persembahkan untuk :

Kedua orang tua saya yang sudah membesarkan, mendidik, membimbing, dan memberikan pendidikan yang layak untuk saya, serta do'a dan dukungannya yang selalu mengiringi setiap langkah saya.

Adik saya yang tercinta dan semua keluarga besar yang senantiasa mendukung dan menyayangi saya.

Almamater tercinta Universitas Islam Negeri Sunan Kalijaga Yogyakarta
Bapak Ibu dosen yang telah memberikan ilmunya serta membimbing saya hingga perkuliahan saya selesai.

Teman-teman yang selalu memberikan dukungan dan motivasi kepada saya.



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

MOTTO

“Menghina Tuhan tidak perlu dengan umpatan dan membakar kitab-Nya. Khawatir besok kamu tak bisa makan saja itu sudah menghina Tuhan.”

~Sujiwo Tejo~

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Simulasi Penanganan Masalah Multikolinearitas dengan Menggunakan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge*” ini dengan baik. Shalawat serta salam semoga tercurahkan kepada Nabi Muhammad SAW yang telah mengantarkan manusia ke jalan kebenaran.

Keberhasilan penulisan skripsi ini tidak lepas dari bimbingan, arahan, dan dukungan dari berbagai pihak, baik berupa pikiran, motivasi, tenaga, maupun do'a. Oleh karena itu, dengan kerendahan hati penulis mengucapkan rasa terima kasih kepada :

1. Prof. Dr. Phil. Al Makin, S. Ag., M. A., selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta
2. Dr. Hj. Khurul Wardati, M. Si., selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta
3. Muchammad Abrori, S. Si., M. Kom., selaku Ketua Program Studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta
4. Mohammad Farhan Qudratullah, S. Si., M. Si., selaku dosen pembimbing skripsi yang telah memberikan ilmu pengetahuan, pengalaman, dan membimbing penulis sehingga mempermudah dalam proses penyusunan skripsi ini.

5. Dr. Epha Diana Supandi, S. Si., M. Sc., dan Dr. Muhammad Wakhid Musthofa, S. Si., M. Si., selaku dosen penguji yang sudah memberikan saran, bimbingan, dan dukungannya sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
6. Bapak/Ibu Dosen dan Staff Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga Yogyakarta atas ilmu, bimbingan, dan pelayanan selama perkuliahan hingga penyusunan skripsi ini selesai.
7. Bapak Karso dan Ibu Evi Mujiati selaku orang tua penulis yang telah memberikan kasih sayang, mendoakan, dan memberi dukungan tiada henti kepada penulis, yang selalu setia menjadi tempat curahan dan merestui setiap langkah penulis.
8. Anindia Giani Garnita selaku adik penulis, Ibu Roisah selaku nenek dari penulis, serta semua keluarga yang telah memberikan kasih sayang, perhatian, semangat, dan tidak hentinya mendoakan keberhasilan penulis.
9. Anatansyah Ayomi Anandari, Sriyanti Mustika Ningrum, Novia Amilatus Solekhah, Zainul Khozin, Ezra Efzrizardi Yusufa, selaku teman-teman yang sudah banyak membantu penulis dalam menyelesaikan skripsi ini.
10. Sahabat-sahabat penulis, Tri Nur Khofifah, Julita Sari, Royhana Devi, Diwanti Panca Satiti, Istinganatun, Riyana Yuni Sulistyawati, Ratih Nur Fatimah, Linaksanan, yang setia mendengarkan curahan hati dan memberikan dukungannya kepada penulis.
11. Teman-teman KKN, Dian Ramadhani, Faridatul Istiqomah, Nur Athifah Apriliyani, Amara Larasakti, yang selalu menjadi teman berkeluh kesah,

bercanda, dan berbagi kerumitan pikiran serta memberkan dukungan dan semangat kepada penulis.

12. Ferninda Dewi Kurniastuti selaku teman dari awal masuk kuliah hingga sekarang, yang selalu setia mendengarkan curahan hati penulis dan tiada henti memberikan dukungannya sehingga penulis dapat menyelesaikan skripsi ini.
13. Anjar Mei Anggoro Wati, Lia Febriyanti, teman-teman satu bimbingan yang selalu berbagi cerita dan pikiran sehingga penulis bersemangat dalam menyelesaikan skripsi ini.
14. Teman-teman prodi Matematika angkatan 2017 yang selalu menemani, memberikan dukungan dan pelajaran berharga selama ini.
15. Kim Namjoon, Kim Seokjin, Min Yoongi, Jung Hoseok, Park Jimin, Kim Taehyung, Jeon Jungkook, selaku orang-orang istimewa yang memotivasi penulis lewat lagu-lagunya, sehingga penulis bersemangat dalam menyelesaikan skripsi ini.
16. Semua pihak yang tidak bisa penulis sebutkan satu persatu, yang memberikan doa dan dukungan kepada penulis dalam menyelesaikan skripsi ini.

Semoga Allah SWT menerima amal kebaikan beliau sekalian dan memberikan balasan dan pahala yang berlipat.

Penulis menyadari bahwa skripsi ini masih banyak kekurangan dan jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran

yang membangun agar penulis dapat membuat karya dengan lebih baik. Penulis berharap semoga skripsi ini dapat memberikan manfaat kepada pembaca.

Yogyakarta, 19 Agustus 2021

Penulis



DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN.....	iv
HALAMAN PERSEMBAHAN	v
MOTTO	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	xi
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvii
DAFTAR LAMPIRAN	xvii
DAFTAR SIMBOL	xviii
ABSTRAK	xx
ABSTRACT	xxii
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Batasan Masalah.....	4
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
1.6 Tinjauan Pustaka	6
1.7 Sistematika Penulisan.....	11
BAB II	13
LANDASAN TEORI.....	13
2.1 Matriks.....	13
2.1.1 Definisi Matriks	13
2.1.2 Penjumlahan Matriks	15
2.1.3 Perkalian Matriks	15
2.1.4 Transpos Matriks	19

2.1.5	Invers Matriks	19
2.1.6	Trace Matriks	20
2.1.7	Nilai Eigen dan Vektor Eigen	21
2.1.8	Kombinasi Linear	21
2.2	Analisis Regresi	22
2.2.1	Definisi	22
2.2.2	Tahapan Dalam Analisis Regresi	23
2.3	Analisis Regresi Linear Berganda	24
2.3.1	Definisi Analisis Regresi Linear Berganda	24
2.3.2	Tujuan Analisis Regresi Linear Berganda	25
2.3.3	Syarat Penggunaan Analisis Regresi Linear Berganda	25
2.3.4	Model Regresi yang Baik	27
2.3.5	Model Regresi Linear Berganda	28
2.3.6	Estimasi Parameter Model Regresi Linear Berganda	29
2.4	Uji Asumsi Analisis Regresi	30
2.4.1	Uji Persyaratan Statistik	30
2.4.2	Uji Asumsi Klasik	33
2.5	Principal Component Analysis (PCA)	37
2.6	<i>Principal Component Regression</i> (PCR)	38
2.7	Regresi <i>Ridge</i>	40
2.8	Pemilihan Model Terbaik	41
BAB III	43
METODOLOGI PENELITIAN	43
3.1	Jenis Penelitian	43
3.2	Metode Pengumpulan Data	44
3.3	Variabel Penelitian	44
3.4	Metode Analisis Data	44
3.5	Pengujian Hipotesis	46
3.6	Penarikan Kesimpulan	47
3.7	Flowchart	48
BAB IV	49
PEMBAHASAN	49
4.1	Multikolinearitas	49
4.1.1	Estimasi Multikolinearitas Sempurna	50

4.1.2	Estimasi Multikolinearitas Tinggi, namun Tidak Sempurna.....	53
4.1.3	Konsekuensi Multikolinearitas	53
4.2	Principal Component Analysis (PCA)	60
4.2.1	Teknik Penggunaan PCA.....	61
4.2.2	Kriteria Pemilihan Komponen Utama	63
4.3	Principal Component Regression (PCR).....	64
4.3.1	Penggunaan <i>Ordinary Least Square</i> (OLS) pada PCR.....	64
4.3.2	Nilai Ekspektasi dan Variansi Penduga Koefisien PCR.....	65
4.3.3	Uji Keberartian Model PCR	66
4.3.4	Uji Signifikansi Model PCR.....	67
4.4	Regresi <i>Ridge</i>	68
4.4.1	Penentuan Konstanta Bias c	69
4.4.2	Estimator Regresi <i>Ridge</i>	69
4.4.3	Hubungan antara Estimator Regresi <i>Ridge</i> dengan <i>Ordinary Least Square</i> (OLS)	72
4.4.4	Sifat-sifat Estimator Regresi <i>Ridge</i>	73
4.4.5	Pendeteksian Multikolinearitas dengan Regresi <i>Ridge</i>	76
BAB V	79
SIMULASI	79
5.1	Deskripsi Data	79
5.2	Pendeteksian Multikolinearitas	81
5.2.1	Uji Simultan F	82
5.2.2	Mendeteksi Multikolinearitas	82
5.3	Penanganan Multikolinearitas dengan Metode <i>Principal Component Regression</i> (PCR)	84
5.4	Penanganan Multikolinearitas dengan Metode Regresi <i>Ridge</i>	96
5.4.1	Pemodelan Regresi <i>Ridge</i>	96
5.4.2	Uji Keberartian Regresi	102
5.5	Pemilihan Model Terbaik	104
BAB VI	107
PENUTUP	107
6.1	Kesimpulan.....	107
6.2	Saran	109
DAFTAR PUSTAKA	110
LAMPIRAN	112



DAFTAR TABEL

Tabel 1.1 Tinjauan Pustaka	7
Tabel 1.2 Perbedaan dengan Penelitian Terdahulu	10
Tabel 4.1 Dampak Peningkatan Multikolinearitas pada Interval Kepercayaan 95%	57
Tabel 4.2 Data Hipotesis pada Y , X_2 , dan X_3	58
Tabel 4.3 Data Hipotesis pada Y , X_2 , dan X_3	59
Tabel 4.4 ANOVA	66
Tabel 5.1 Data Ulangan ke-1 yang Telah Dibangkitkan dari <i>Software R</i>	79
Tabel 5.2 Estimasi Parameter Regresi Kuadrat Terkecil Ulangan ke-1	81
Tabel 5.3 Nilai VIF	83
Tabel 5.4 Tabel Komponen Utama Ulangan ke-1	85
Tabel 5.5 Tabel Harga Skor Komponen Utama Ulangan ke-1	86
Tabel 5.6 Model PCR 30 Kali Ulangan	94
Tabel 5.7 Nilai VIF $\hat{\beta}(c)$ dengan Berbagai Nilai c	96
Tabel 5.8 Nilai $\hat{\beta}(c)$ dengan Berbagai Nilai c	98
Tabel 5.9 Model Regresi <i>Ridge</i> 30 Kali Ulangan.....	100
Tabel 5.10 ANOVA <i>Ridge</i>	103
Tabel 5.11 Nilai MSE.....	104

DAFTAR GAMBAR

Gambar 5.1 <i>Ridge Trace</i>	99
Gambar 5.2 VIF Plot	99



DAFTAR LAMPIRAN

Lampiran 1 Program Membangkitkan Data dengan 30 kali ulangan dengan <i>Software R</i>	113
Lampiran 2 Data yang Ulangan Ke-1 yang Telah Dibangkitkan dari <i>Software R</i>	113
Lampiran 3 Program dan Output Model Regresi	115
Lampiran 4 Program dan Output ANOVA	116
Lampiran 5 Program dan Output Pendeteksian Multikolinearitas dengan Nilai VIF	116
Lampiran 6 Program dan Output <i>Principal Component Analysis</i> (PCA)	117
Lampiran 7 Output NCSS Regresi <i>Ridge</i>	118
Lampiran 8 Perhitungan Nilai <i>Mean Squared Error</i> (MSE).....	124
Lampiran 9 Tabel Distribusi F	126

DAFTAR SIMBOL

I	: Matriks identitas
A^{-1}	: Invers matriks A
w	: Kombinasi linear
Y	: variabel dependen
X	: variabel independen
α	: <i>Intercept</i> (titik potong) kurva terhadap sumbu Y
β	: Kemiringan (<i>slope</i>) kurva linear
ε	: Variabel pengganggu (residual)
β	: Koefisien variabel independen
$\hat{\beta}$: Parameter
Var	: Variansi
E	: Estimator
n	: Ukuran sampel
k	: Jumlah variabel independen
β_p	: Koefisien regresi pada variabel X_p
1	: Vektor yang elemennya adalah 1 berukuran $n \times 1$
α_k	: vektor koefisien komponen utama berukuran $k \times 1$
\hat{Y}_i	: variabel tak bebas dugaan
R^2	: koefisien determinasi
Cov	: Kovariansi

$r_{2\ 3}$: koefisien korelasi antara X_2 dan X_3

A : matriks yang bertransformasi terhadap variabel asal (X), sehingga diperoleh vektor komponen Y .

λ_k : Nilai *eigen* dari ρ

L : Pengali *Lagrange*

Z : Variabel yang distandarisasi

W : Komponen utama



**SIMULASI PENANGANAN MASALAH MULTIKOLINEARITAS
DENGAN MENGGUNAKAN METODE *PRINCIPAL COMPONENT
REGRESSION (PCR)* DAN METODE REGRESI *RIDGE***

Oleh: Hannifa Irmajihan Nabilla

ABSTRAK

Analisis regresi merupakan analisis yang digunakan untuk mengetahui pengaruh antara variabel bebas dan variabel terikat. Salah satu syarat yang harus dipenuhi dalam asumsi klasik adalah tidak terdapat multikolinearitas di antara variabel. Multikolinearitas merupakan kondisi dimana dua atau lebih variabel bebas memiliki hubungan linear sempurna atau hampir sempurna (kolinear). Dampak dari multikolinearitas yaitu meskipun bersifat BLUE (*Best Linear Unbiased Estimator*) akan tetapi estimator *Ordinary Least Square (OLS)* memiliki variansi dan kovariansi yang besar sehingga membuat estimator yang akurat menjadi sulit, interval kepercayaan cenderung lebih luas, rasio t dari satu atau lebih koefisien cenderung tidak signifikan, meskipun nilai R^2 tinggi namun hanya sedikit rasio t yang signifikan, dan estimator OLS beserta standar errornya bersifat sensitif akan perubahan kecil pada data. Untuk mengatasi masalah multikolinearitas dalam penelitian ini penulis menggunakan dua metode yaitu metode *Principal Component Regression (PCR)* dan metode Regresi *Ridge*. PCR merupakan metode yang menggabungkan antara *Principal Component Analysis (PCA)* dan Analisis Regresi. Sedangkan metode Regresi *Ridge* merupakan pengembangan dari OLS yang menghasilkan penduga bersifat bias dari koefisien regresi.

Penelitian ini bertujuan untuk mengetahui metode terbaik antara metode PCR dan metode Regresi *Ridge* dengan kriteria *Mean Squared Error (MSE)* yang memiliki nilai minimum. Nilai MSE diperoleh dengan cara mengkuadratkan selisih antara nilai ramalan dengan nilai aktual. Metode PCR dan metode Regresi *Ridge* diterapkan pada data simulasi yang dibangkitkan penulis dari *software R* dengan 30 kali ulangan. Variabel yang digunakan yaitu variabel dependen Y dengan melibatkan lima variabel independen (X_1, X_2, X_3, X_4 dan X_5), dimana variabel X_1 dan X_5 mengandung unsur multikolinearitas.

Setelah dilakukan analisis data, diperoleh rata-rata nilai MSE dari model PCR sebesar 2487,69, dan rata-rata nilai MSE dari model Regresi *Ridge* sebesar 1036,92. Dengan demikian pada simulasi ini metode Regresi *Ridge* lebih efektif digunakan untuk mengatasi masalah multikolinearitas dibandingkan metode PCR.

Kata Kunci: *Mean Squared Error, Multikolinearitas, Principal Component Analysis, Principal Component Regression, Regresi Ridge.*

**SIMULATION OF HANDLING MULTICOLLINEARITY PROBLEMS
USING PRINCIPAL COMPONENT REGRESSION (PCR) METHODS
AND RIDGE REGRESSION METHODS**

By: Hannifa Irmajihan Nabilla

ABSTRACT

Regression analysis is an analysis used to determine the effect between the independent variable and the dependent variable. One of the conditions that must be met in the classical assumption is that there is no multicollinearity between variables. Multicollinearity is a condition in which two or more independent variables have a perfect or almost perfect (collinear) linear relationship. The impact of multicollinearity is that although it is a BLUE (Best Linear Unbiased Estimator) but the Ordinary Least Square (OLS) estimator has a large variance and covariance so that it makes an accurate estimator difficult, the confidence interval tends to be wider, the t ratio of one or more coefficients tends to not significant, even though the R^2 value is high, the t ratio is only slightly significant, and the OLS estimator and its standard error are sensitive to small changes in the data. To overcome the problem of multicollinearity in this study the author uses two methods, namely the Principal Component Regression (PCR) method and the Ridge Regression method. PCR is a method that combines Principal Component Analysis (PCA) and Regression Analysis. While the Ridge Regression method is the development of the Ordinary Least Square (OLS) which produces a biased estimator of the regression coefficients.

This study aims to determine the best method between the PCR method and the Ridge Regression method with the Mean Squared Error (MSE) criterion which has a minimum value. The MSE value is obtained by squaring the difference between the forecast value and the actual value. The PCR method and the Ridge Regression method were applied to the simulation data generated by the author from the R software with 30 replications. The variable used was the dependent variable Y involving five independent variables ($X_1, X_2, X_3, X_4,$ and X_5), where the variable X_1 and X_5 contained elements of multicollinearity.

After analyzing the data, the average MSE value from the PCR model was 2487,69, and the average MSE value from the Ridge Regression model was 1036,92. Thus, in this simulation the Ridge Regression method is more effectively used to overcome the problem of multicollinearity than the PCR method.

Keywords: Mean Squared Error, Multicollinearity, Principal Component Analysis, Principal Component Regression, Ridge Regression.

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Statistika merupakan ilmu pengetahuan yang memuat kegiatan koleksi data, presentasi data, analisis data, dan interpretasi data (Qudratullah, 2013). Sampai dengan abad ke-19, para matematikawan seperti Pascal, Bernaulli, De Moivre, Laplace, Gauss, dan yang lainnya menggunakan statistika untuk mendapatkan informasi mengenai pendapatan, penduduk, dan bidang yang berkaitan dengan pemerintahan atau kerajaan (Sugiyono, 2013). Terdapat beberapa macam jenis statistika yang digolongkan berdasarkan orientasi pembahasan ataupun berdasarkan tujuan analisisnya. Berdasarkan orientasi pembahasan, statistika dibagi menjadi dua, yaitu statistika matematik dan statistika terapan. Lalu berdasarkan tujuan atau tahap analisisnya, statistika dibagi menjadi statistika deskriptif dan statistika inferensia. Selain pengelompokan tadi, statistika juga biasa menggunakan data distribusi populasi, yang dibedakan menjadi statistika parametrik dan nonparametrik. Di dalam statistika ada tiga jenis data, yakni data *cross section*, data *time series*, dan data panel. Data *cross section* biasanya dianalisis menggunakan analisis regresi (Qudratullah, 2013).

Analisis regresi adalah salah satu analisis yang paling terkenal di kalangan peneliti. Analisis regresi banyak digunakan di tingkat pendidikan S1, S2, maupun S3. Karena analisis ini sangatlah populer maka dapat disimpulkan bahwa selama ini setiap kejadian adalah terkait dan saling mempengaruhi (Ariyanto, 2005). Di dalam analisis regresi kita mempelajari bagaimana cara membentuk suatu model

fungsional dari data agar dapat menjelaskan atau meramalkan suatu peristiwa alami atas dasar peristiwa yang lain. Analisis regresi juga dinyatakan sebagai suatu analisis mengenai hubungan antara dua variabel atau lebih yang pada umumnya disajikan dalam persamaan matematika. Analisis regresi digunakan untuk mengetahui pengaruh antara variabel bebas dan variabel terikat. Variabel bebas merupakan variabel yang mempengaruhi, disebut sebagai variabel independen atau prediktor, dan disimbolkan dengan X . Sedangkan variabel terikat merupakan variabel yang dipengaruhi, disebut sebagai variabel dependen, dan disimbolkan dengan Y (Ariyanto, 2005).

Menurut Quadratullah (2013) analisis regresi dibagi menjadi tiga, yaitu analisis regresi parametrik, analisis regresi nonparametrik, dan analisis regresi semiparametrik (perpaduan antara regresi parametrik dan regresi nonparametrik). Perbedaan utama antara regresi parametrik dan regresi nonparametrik adalah analisis regresi parametrik memerlukan asumsi-asumsi baik bentuk fungsional maupun distribusi residualnya. Tetapi dapat pula digunakan analisis regresi nonparametrik apabila asumsi-asumsi tidak terpenuhi.

Bentuk paling sederhana dalam analisis regresi adalah regresi linear sederhana. Analisis regresi linear sederhana hanya melibatkan dua variabel, yaitu satu variabel terikat (dependen) dan satu variabel bebas (independen). Sedangkan analisis regresi yang melibatkan dua atau lebih variabel independen disebut analisis regresi linear berganda (Ifadah, 2011).

Di dalam statistika sebuah model dikatakan baik apabila telah memenuhi asumsi klasik, yaitu tidak mengandung autokorelasi, heterokedastisitas, dan

multikolinearitas. Maka proses kendali terhadap model perlu dilakukan untuk mengetahui terpenuhi atau tidaknya asumsi-asumsi tersebut.

Salah satu dari ketiga syarat yang harus dipenuhi dalam asumsi klasik adalah tidak terdapat multikolinearitas di antara variabel. Ditemukannya multikolinearitas ketika menentukan model regresi populasi. Terdapat kemungkinan bahwa sampel tertentu, beberapa atau semua variabel X memiliki hubungan linear sempurna atau hampir sempurna (kolinear).

Ada beberapa langkah yang digunakan dalam mengatasi masalah multikolinearitas. Apabila pemilihan variabel diperbolehkan dan tidak mengubah teori yang sudah ada maka cara yang termudah dalam mengatasi masalah multikolinearitas adalah dengan mengeluarkan salah satu variabel independen yang tidak penting dalam model sehingga diperoleh estimator dengan varian lebih kecil. Akan tetapi tidak semua permasalahan yang di dalamnya terjadi multikolinearitas dapat diatasi menggunakan metode tersebut karena dapat mempengaruhi variabel dependen. Maka dari itu diperlukan metode lain yang dalam prosesnya tidak mengeluarkan variabel independen dalam model regresi dan metode estimasi lain yang dapat menghasilkan parameter dengan variansi yang lebih kecil. Metode lain yang akan digunakan di dalam penelitian ini adalah Metode *Principal Component Analysis* (PCA) dan Metode Regresi *Ridge*.

Dengan digunakannya Metode *Principal Component Analysis* (PCA) akan dihasilkan variabel-variabel baru yang merupakan kombinasi linear dari variabel-variabel bebas asal. Variabel baru ini disebut sebagai komponen utama yang kemudian diregresikan dengan variabel tak bebas. Analisis yang menggabungkan

antara PCA dan analisis regresi disebut dengan Metode *Principal Component Regression* (PCR). Sedangkan Metode Regresi *Ridge* menghasilkan taksiran koefisien regresi dengan varian yang lebih kecil, namun taksiran koefisien regresinya bersifat bias (Ifadah, 2011).

1.2 Batasan Masalah

Dilihat dari latar belakang yang sudah dipaparkan, untuk menghindari terdapatnya ruang lingkup masalah yang terlalu luas dalam penelitian ini, maka peneliti memberikan batasan masalah sebagai berikut :

1. Analisis yang digunakan dalam penelitian ini adalah Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge*.
2. Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* digunakan dalam mengatasi masalah multikolinearitas pada data.
3. Alat analisis yang digunakan dalam penelitian ini adalah *software* komputer yaitu program R Studio, NCSS, dan Microsoft Excel.

1.3 Rumusan Masalah

Berdasarkan batasan masalah yang telah peneliti pilih, maka dapat dirumuskan permasalahan dalam penelitian ini sebagai berikut :

1. Bagaimana langkah-langkah Metode *Principal Component Regression* (PCR) dalam mengatasi masalah multikolinearitas?
2. Bagaimana langkah-langkah Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas?
3. Manakah antara Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* yang lebih efektif digunakan untuk mengatasi masalah

multikolinearitas dengan menggunakan data simulasi yang dibangkitkan dari *software R*?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk mengetahui langkah-langkah Metode *Principal Component Regression* (PCR) dalam mengatasi masalah multikolinearitas.
2. Untuk mengetahui langkah-langkah Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas.
3. Untuk mengetahui mengetahui metode yang lebih efektif antara Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas dengan menggunakan data simulasi yang dibangkitkan dari *software R*.

1.5 Manfaat Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka manfaat dari penelitian ini adalah sebagai berikut :

1. Bagi Peneliti
 - a. Sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Matematika (S1).
 - b. Menambah pengetahuan terkait dengan metode analisis data dalam Statistika dan pengaplikasiannya.

- c. Mengetahui dan memahami penggunaan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* dalam hal mengatasi masalah multikolinearitas pada Analisis Regresi Linear Berganda.
- d. Melatih ketelitian dan keuletan dalam mengatasi kendala yang terjadi saat menganalisis suatu masalah dalam Statistika.

2. Bagi Program Studi Matematika

Besar harapan penulis agar penelitian ini dapat dimanfaatkan, dipelajari, dan dipahami dengan baik sebagai bahan acuan atau referensi yang membantu dalam penelitian Program Studi Matematika selanjutnya. Penelitian ini juga memberikan tambahan ilmu dan wawasan baru mengenai penggunaan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* dalam hal mengatasi masalah multikolinearitas pada Analisis Regresi Linear Berganda.

3. Bagi Pembaca

Menjadi bahan acuan sekaligus inspirasi bagi pembaca yang ingin mengembangkan atau menghadapi permasalahan yang berkaitan dengan penelitian ini.

1.6 Tinjauan Pustaka

Pada penelitian ini, peneliti menggunakan beberapa tinjauan pustaka seperti buku, jurnal, dan skripsi yang berkaitan. Berikut merupakan sumber referensi yang digunakan peneliti dalam membantu menyelesaikan penelitian:

Tabel 1.1 Tinjauan Pustaka

Nama Peneliti	Judul	Studi Kasus	Software
T. L. Wasilane, M. W. Talakua, dan Y. A. Lesnussa	Model Regresi <i>Ridge</i> untuk Mengatasi Model Regresi Linear Berganda yang Mengandung Multikolinearitas	Data Pertumbuhan Bayi di Kelurahan Namaelo RT 001, Kota Masohi)	SPSS dan NCSS
Agustifa Zea Tazliqoh, Rita Rahmawati, Diah Safitri	Perbandingan Regresi Komponen Utama dengan Regresi <i>Ridge</i> pada Analisis Faktor- Faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah	Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah	SPSS dan MATLAB R2010a
Wenty Resti Anggraeni, Naomi Nessyana Debataraja, Setyo Wira Rizki	Estimasi Parameter Regresi <i>Ridge</i> untuk Mengatasi Multikolinearitas	Data tingkat pengangguran terbuka, jumlah penduduk miskin, tingkat partisipasi	NCSS

		angkatan kerja dari provinsi di Indonesia	
Drajat Indra Purnama	Analisis Komponen Utama pada Data Potensi Kecamatan di Kota Palu Sebelum Bencana Gempa Bumi dan Tsunami 28 September 2018	Data Potensi Kecamatan di Kota Palu Sebelum Bencana Gempa Bumi dan Tsunami 28 September 2018	R seri 3.5.1
Tatik Widiharah	Penanganan Multikolinearitas (Kekolinearan Ganda) dengan Analisis Regresi Komponen Utama	Penjualan suatu produk dalam ribuan dolar dengan potensi wilayah, biaya promosi dalam puluhan dolar, banyaknya merek saingan dan banyaknya <i>active accounts</i> dari 13 wilayah penjualan	Minitab 11.12

G. L. Marcus, H. J. Wattimanela, Y. A. Lesnussa	Analisis Regresi Komponen Utama untuk Mengatasi Masalah Multikolinearitas dalam Analisis Regresi Linear Berganda	Curah hujan di Kota Ambon tahun 2010	
Hannifa Irmajihan Nabilla (17106010042)	Simulasi Penanganan Masalah Multikolinearitas dengan Menggunakan Metode <i>Principal Component Regression</i> (PCR) dan Metode Regresi <i>Ridge</i>	Simulasi menggunakan data yang dibangkitkan dari <i>software</i> R	R Studio, NCSS, Microsoft Excel

Persamaan penelitian ini dengan enam penelitian sebelumnya yaitu membahas penyelesaian multikolinearitas pada model regresi. Sedangkan perbedaan penelitian ini dengan penelitian sebelumnya adalah sebagai berikut:

Tabel 1.2 Perbedaan dengan Penelitian Terdahulu

Nama Peneliti	Metode Penyelesaian	Studi Kasus/Simulasi
T. L. Wasilane, M. W. Talakua, dan Y. A. Lesnussa	Metode Regresi <i>Ridge</i>	Studi kasus
Agustifa Zea Tazliqoh, Rita Rahmawati, Diah Safitri	Metode Regresi Komponen Utama dan Regresi <i>Ridge</i>	Studi kasus
Wenty Resti Anggraeni, Naomi Nesyana Debataraja, Setyo Wira Rizki	Metode Regresi <i>Ridge</i>	Studi kasus
Drajat Indra Purnama	Metode Analisis Komponen Utama	Studi kasus
Tatik Widiharah	Metode Analisis Regresi Komponen Utama	Studi kasus
G. L. Marcus, H. J. Wattimanela, Y. A. Lesnussa	Analisis Regresi Komponen Utama	Studi kasus

Hannifa Irmajihan	Metode <i>Principal Component</i>	
Nabilla (17106010042)	<i>Regression</i> (PCR) dan metode Regresi <i>Ridge</i>	Simulasi

1.7 Sistematika Penulisan

Sistematika penulisan dibuat bertujuan untuk memudahkan pembaca dalam memahami isi dari penelitian ini secara sederhana, rinci, runtut, jelas dan tidak membingungkan. Adapun sistematika penulisan dalam penelitian ini adalah sebagai berikut :

BAB 1 : PENDAHULUAN

Pada bab ini berisi tentang hal-hal yang melatar belakangi peneliti dalam melakukan penelitian, batasan masalah yang diangkat, rumusan masalah yang mendasari penelitian, tujuan dan manfaat dari penelitian, tinjauan pustaka yang menjadi alat bantu peneliti dalam menyelesaikan penelitian, serta sistematika penulisan sebagai gambaran sederhana dari penelitian ini.

BAB II : LANDASAN TEORI

Pada bab ini berisi teori-teori yang mendasari peneliti dalam melakukan penelitian. Teori digunakan sebagai penunjang dan penguat pembahasan terkait penelitian yang dilakukan yaitu perbandingan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas.

BAB III : METODOLOGI PENELITIAN

Pada bab ini berisi jenis penelitian , sumber data, teknik pengolahan data, metode pengumpulan data, metode analisis data, langkah-langkah analisis data yang digunakan oleh peneliti dalam menyelesaikan penelitian ini.

BAB IV : PEMBAHASAN

Pada bab ini berisi pembahasan mengenai metode analisis yang digunakan peneliti dalam menyelesaikan penelitian, yaitu Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas.

BAB V : SIMULASI

Pada bab ini berisi penerapan metode yang dibahas pada bab sebelumnya dengan melakukan simulasi dan data yang dibangkitkan dari *software R*.

BAB VI : PENUTUP

Pada bab ini berisi penarikan kesimpulan dari analisis yang telah dilakukan pada bab sebelumnya. Selain kesimpulan, pada bab ini juga terdapat kritik dan saran untuk penelitian selanjutnya yang masih berhubungan erat dengan penelitian ini.

BAB VI

PENUTUP

Metode *Principal Component Regression* (PCR) dan model Regresi *Ridge* telah dijelaskan dalam pembahasan dan dilakukan simulasi untuk mengatasi masalah multikolinearitas pada data yang dibangkitkan dari *software* R. Sehingga dari berbagai analisis data yang telah dilakukan diperoleh kesimpulan dan saran sebagai berikut.

6.1 Kesimpulan

Berdasarkan rumusan masalah dan hasil penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut.

1. Berikut langkah-langkah Metode *Principal Component Regression* (PCR) dalam mengatasi masalah multikolinearitas:
 - a. Melakukan *Principal Component Analysis* (PCA)
 - b. Mencari nilai eigen dan vektor eigen
 - c. Menghitung skor komponen utama
 - d. Menentukan jumlah komponen utama yang akan digunakan
 - e. Membentuk model *Principal Component Regression* (PCR) menggunakan Metode Ordinary Least Square (OLS) biasa.
2. Berikut langkah-langkah Metode Regresi *Ridge* dalam mengatasi masalah multikolinearitas:
 - a. Mencari nilai VIF $\hat{\beta}(c)$ dengan Berbagai Nilai c
 - b. Mencari nilai koefisien estimator parameter $\hat{\beta}(c)$ dengan berbagai kemungkinan tetapan bias c

- c. Membentuk gambar *Ridge Trace* dari nilai koefisien estimator parameter $\hat{\beta}(c)$
 - d. Membentuk model Regresi *Ridge*
 - e. Melakukan uji keberartian regresi.
3. Data yang digunakan pada penelitian ini merupakan data yang dibangkitkan dari *software* R dengan 30 kali ulangan. Variabel yang digunakan yaitu variabel dependen Y dengan lima variabel independen (X_1, X_2, X_3, X_4 dan X_5), dimana variabel X_1 dan X_5 mengandung unsur multikolineritas. Pada penelitian ini permasalahan multikolinearitas diatasi menggunakan Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge*. Lalu dilakukan pemilihan model terbaik antara Metode *Principal Component Regression* (PCR) dan Metode Regresi *Ridge* untuk mengetahui metode yang lebih efektif digunakan untuk mengatasi masalah multikolinearitas dengan kriteria nilai *Mean Squared Error* (MSE) minimum. Berdasarkan teori yang sudah ada bahwa semakin kecil nilai MSE maka semakin akurat nilai suatu ramalan. Setelah dilakukan perhitungan nilai MSE kedua model dari 30 kali pengulangan, diperoleh rata-rata nilai MSE model PCR adalah 2487,69 dan nilai rata-rata nilai MSE model Regresi *Ridge* adalah 1036,92. Terlihat bahwa rata-rata nilai MSE model Regresi *Ridge* lebih kecil dibandingkan rata-rata nilai MSE model PCR. Sehingga dapat disimpulkan bahwa Regresi *Ridge* lebih efektif digunakan untuk mengatasi masalah multikolinearitas pada data yang dibangkitkan dari *software* R dengan 30 kali pengulangan.

6.2 Saran

Untuk memperluas pemahaman mengenai penelitian ini, ada beberapa saran yang penulis sampaikan sebagai berikut.

1. Apabila terdapat masalah multikolinearitas pada data, maka perlu dilakukan penanganan untuk menghilangkan multikolinearitas tersebut. Karena multikolinearitas dapat memberikan dampak buruk pada data.
2. Ada beberapa metode lain yang dapat digunakan untuk mengatasi masalah multikolinearitas, seperti metode *Partial Least Square*, Analisis Faktor, *Latent Root Regression*, dan lain-lain.
3. Ada beberapa metode lain yang dapat digunakan untuk menentukan model terbaik, seperti metode *Akaike's Information Criterion (AIC)*, *Bayesian Information Criterion (BIC)*, dan lain-lain.

DAFTAR PUSTAKA

- Anggreni, W. R, dkk. 2018. *Estimasi Parameter Regresi Ridge untuk Mengatasi Multikolinearitas*. Buletin Ilmiah Mat, Stat, dan Terapannya (Bimaster). 7(4): 295-302.
- Anton, Howard. 1991. *Aljabar Linear Elementer*. Jakarta: Penerbit Erlangga.
- Budiyono. 2014. *Statistika Untuk Penelitian*, Edisi Kedua, Solo: Sebelas Maret University Press.
- Ghozali, Imam. 2013. *Statistik Nonparametrik*. Semarang: Badan Penerbit UNDIP.
- Gujarati, D. N dan Dawn, C. P. 2010. *Dasar-dasar Ekonometrika*, Edisi Kelima, Jakarta: Penerbit Salemba Empat
- Ifadah, Ana. 2011. Analisis Metode *Principal Component Analysis* (Komponen Utama) dan Regresi *Ridge* dalam Mengatasi Dampak Multikolinearitas dalam Analisis Regresi Linear Berganda. *Skripsi*. Program Studi Matematika Universitas Negeri Semarang. Semarang.
- Marcus, G. L, dkk. 2012. *Analisis Regresi Komponen Utama untuk Mengatasi Masalah Multikolinearitas dalam Analisis Regresi Berganda*: Jurusan Matematika FMIPA UNPATTI. Jurnal Barekeng. 6(1): 31-40.
- Ohyver, Margaretha. 2011. *Metode Regresi Ridge untuk Mengatasi Kasus Multikolinear*: Jurusan Matematika Fakultas Sains dan Teknologi Binus University. Comtect 2(1): 451-457.
- Pratiwi, Novi Bekti. 2016. Perbandingan Regresi Komponen Utama dengan Regresi *Ridge* untuk Mengatasi Masalah Multikolinearitas. *Skripsi*. Program Studi Matematika Universitas Negeri Semarang. Semarang.
- Purnama, Drajat Indra. 2019. *Analisis Komponen Utama pada Data Potensi Kecamatan di Kota Palu Sebelum Gempa Bumi dan Tsunami 28 September 2018*: Prodi Magister Statistika Terapan Universitas Padjadjaran. Jurnal Matematika, Statistika, & Komputasi. 16(1): 25-32.
- Putri, Agriska Prenadita. 2011. Penggunaan Metode *Ridge Trace* dan *Variance Inflation Factors* (VIF) pada Regresi *Ridge*. *Skripsi*. Jurusan Pendidikan Matematika Universitas Negeri Yogyakarta. Yogyakarta.

- Rohmah, Ani. 2013. Perbandingan Metode *Ridge* dengan Metode *Principal Component Analysis* (PCA) dalam Mengatasi Masalah Multikolinearitas. *Skripsi*. Program Studi Matematika UIN Sunan Kalijaga. Yogyakarta.
- Quadratullah, Mohammad Farhan. 2013. *Analisis Regresi Terapan : Teori, Contoh Kasus, dan Aplikasi dengan SPSS*, Yogyakarta: Penerbit ANDI Yogyakarta.
- Tazliqoh, A. Z, dkk. 2015. *Perbandingan Regresi Komponen Utama dengan Regresi Ridge pada Analisis Faktor-faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah*: Jurusan FSM Universitas Diponegoro. *Jurnal Gaussian*. 4(1): 1-10.
- Wahyuningsih, Tri. 2018. Mengatasi Dampak Multikolinearitas dalam Analisis Regresi Linear Berganda dengan Metode Regresi *Ridge* dan Metode *Principal Component Analysis* (PCA) (Studi Kasus Hasil Penjualan Ana Bakery Tahun 2017 di Yogyakarta). *Skripsi*. Program Studi Matematika UIN Sunan Kalijaga. Yogyakarta.
- Wasilaine, T. L, dkk. 2014. *Model Regresi Ridge untuk Model Regresi Linear Berganda yang Mengandung Multikolinearitas*: Jurusan Matematika FMIPA UNPATTI. *Jurnal Barekeng*. 8(1): 31-37.
- Widiharih, Tatik. 2001. *Penanganan Multikolinearitas (Kekolinearan Ganda) dengan Analisis Regresi Komponen Utama*: Jurusan Matematika FMIPA UNDIP. *Jurnal Matematika dan Komputer*. 4(2): 71-81.
- Sugiyono. 2013. *Statistika Untuk Penelitian*, Bandung: Penerbit ALFABETA.