

**PENERAPAN METODE REGRESI *RIDGE* UNTUK MENGATASI
MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS
PEMBANGUNAN MANUSIA DI PROVINSI JAWA TENGAH**

TUGAS AKHIR



Disusun Oleh:

M. Zamroni Rosyadi

14 611 118

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2018**

**PENERAPAN METODE REGRESI *RIDGE* UNTUK MENGATASI
MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS
PEMBANGUNAN MANUSIA DI PROVINSI JAWA TENGAH**

TUGAS AKHIR

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Jurusan
Statistika**



Disusun Oleh:

M. Zamroni Rosyadi

14 611 118

**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
2018**

HALAMAN PERSETUJUAN PEMBIMBING
TUGAS AKHIR

Judul : Penerapan Metode Regresi *Ridge* Untuk Mengatasi Masalah
Multikolinearitas Pada Kasus Indeks Pembangunan Manusia di
Provinsi Jawa Tengah

Nama Mahasiswa : M Zamroni Rosyadi

Nomor Mahasiswa: 14 611 118

TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK
DIUJIKAN

Yogyakarta, Maret 2018

Pembimbing



(Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.)

HALAMAN PENGESAHAN
TUGAS AKHIR

PENERAPAN METODE REGRESI *RIDGE* UNTUK MENGATASI
MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS
PEMBANGUNAN MANUSIA DI PROVINSI JAWA TENGAH

Nama Mahasiswa : M Zamroni Rosyadi

Nomor Mahasiswa : 14 611 118

TUGAS AKHIR INI TELAH DIUJIKAN
MARET 2018

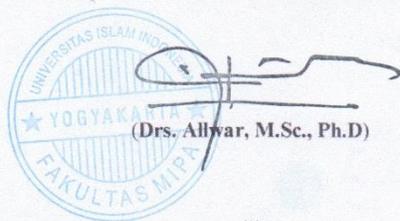
Nama Penguji

Tanda Tangan

1. Annisa Uswatun Khasanah, ST.,
M.Sc.
2. Dr. Edy Widodo, M.Si.
3. Muhammad Hasan Sidiq
Kurniawan, S.Si., M.Sc.

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



KATA PENGANTAR



Assalamu'alaikum Wr.Wb.

Puji syukur penyusun panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya berupa keimanan, kekuatan, kesabaran, kelancaran serta keselamatan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Metode Regresi *Ridge* Untuk Mengatasi Masalah Multikolinearitas Pada Kasus Indeks Pembangunan Manusia di Provinsi Jawa Tengah”. Shalawat serta salam tercurah kepada junjungan kita Nabi Muhammad SAW beserta keluarga dan para pengikut-pengikutnya.

Penulisan skripsi ini disusun sebagai salah satu persyaratan untuk memperoleh gelar Sarjana Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Islam Indonesia. Selama penyusunan skripsi ini, penyusun telah banyak mendapat bimbingan dan bantuan dari berbagai pihak. Untuk itu pada kesempatan ini penyusun bermaksud menyampaikan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Drs. Allwar, M.Sc., Ph.D., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia.
2. Bapak Dr. RB. Fajriya Hakim, S.Si., M.Si., selaku Ketua Jurusan Statistika beserta seluruh jajarannya.
3. Bapak Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc. yang sangat berjasa dalam penyelesaian skripsi ini dan selalu memberi bimbingan selama penulisan skripsi ini.
4. Ibu Kariyam, S.Si., M.Si., selaku dosen pembimbing dalam menyusun tugas akhir ini yang telah memberikan bimbingan dan arahnya.
5. Seluruh Dosen dan Staff Program Studi Statistika Universitas Islam Indonesia yang selalu berbagi ilmu baik dalam bidang akademik maupun non akademik.
6. Dewan Penguji yang telah memberikan saran dalam penulisan skripsi ini.

7. Bapak, Ibu, dan Keluarga Besar yang selalu mendoakan yang terbaik untuk saya.
8. Teman-teman satu bimbingan tugas akhir (bimbingan bapak hasan) Panji, Irsyad, Ulin, Rima, Rati, Tista, Yusi, Ajeng, Indah, Dhea, Ina, Elisa, Marisa dan Nilam yang selalu berbagi ilmu dan berbagi cerita serta pengalaman.
9. Teman-teman saya : Fahri, Yadin, Dharma, Fandi, Feri, Bana, Shodiq, Kholis, Odon, yang sudah banyak memberikan semangat dan bantuan dalam memulai dan mengakhiri tugas kerja praktek ini.
10. Teman-teman KKN unit 405 Kecamatan Karangdowo, Dusun Ringin Putih, Dandi, Dio, Rico, Herni, Rohini, Rahma, Janet dan Siti, suka dan duka yang telah dilalui bersama tidak akan pernah terlupakan.
11. Teman-teman Statistika 2014 yang sudah banyak memberikan semangat dan dukungan selama penulisan skripsi ini.
12. Semua pihak yang tidak dapat penyusun sebutkan satu per satu, terima kasih.

Penyusun menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna, oleh karena itu segala kritik dan saran yang sifatnya membangun selalu penyusun harapkan. Semoga skripsi ini dapat bermanfaat bagi penyusun khususnya dan bagi semua yang membutuhkan umumnya. Akhir kata, semoga Allah SWT selalu melimpahkan rahmat serta hidayah-Nya kepada kita semua, Amin amin ya robbal 'alamiin

Wassalamu'alaikum, Wr.Wb .

Yogyakarta, 19 Februari 2018

M. Zamroni Rosyadi

DAFTAR ISI

| | |
|-----------------------------------------------------------------------------|------|
| HALAMAN JUDUL..... | i |
| HALAMAN PERSETUJUAN PEMBIMBING | ii |
| HALAMAN PENGESAHAN TUGAS AKHIR..... | iii |
| KATA PENGANTAR | iv |
| DAFTAR ISI..... | vi |
| DAFTAR TABEL..... | viii |
| DAFTAR GAMBAR | ix |
| DAFTAR LAMPIRAN..... | x |
| DAFTAR ISTILAH | xi |
| PERNYATAAN..... | xii |
| INTISARI..... | xiii |
| ABSTRACT | xiv |
| BAB I. PENDAHULUAN | |
| 1.1. Latar Belakang | 1 |
| 1.2. Rumusan Masalah | 5 |
| 1.3. Jenis Penelitian dan Metode Analisis..... | 5 |
| 1.4. Tujuan Penelitian | 5 |
| 1.5. Batasan Masalah..... | 6 |
| 1.6. Manfaat Penelitian | 6 |
| BAB II. TINJAUAN PUSTAKA | |
| 2.1. Penelitian Sebelumnya | 7 |
| BAB III. LANDASAN TEORI | |
| 3.1. Model Regresi Linear Berganda | 10 |
| 3.2. Metode Kuadrat Terkecil | 10 |
| 3.3. Sifat-Sifat Estimator Kuadrat Terkecil | 14 |
| 3.4. Ukuran Pemusatan dan Penskalaan (<i>Centering and Scaling</i>) | 14 |

| | |
|--------------------------------------------------------------------|----|
| 3.5. Matriks Korelasi..... | 16 |
| 3.6. Multikolinearitas | 18 |
| 3.7. Regresi <i>Ridge</i> | 19 |
| 3.8. Mendeteksi Multikolinearitas dengan Metode <i>Ridge</i> | 21 |
| 3.9. Pengujian Signifikan Parameter Model | 23 |
| 3.10. <i>Mean of Square Error</i> (MSE) | 24 |
| 3.11. Koefisien Determinasi..... | 25 |
| 3.12. IPM..... | 25 |
| BAB IV. METODE PENELITIAN | |
| 4.1. Data | 27 |
| 4.2. Tempat dan Waktu Penelitian | 27 |
| 4.3. Variabel Penelitian | 27 |
| 4.4. Metode Pengumpulan Data | 28 |
| 4.5. Alat dan Cara Organisir Data..... | 29 |
| BAB V. PEMBAHASAN | |
| 5.1. Analisis Deskriptif | 31 |
| 5.2. Analisis Metode Kuadrat Terkecil | 35 |
| 5.3. Uji <i>Overall</i> dan Parsial..... | 38 |
| 5.4. Uji Multikolinearitas | 40 |
| 5.5. Regresi <i>Ridge</i> | 41 |
| BAB VI. PENUTUP | |
| 6.1. Kesimpulan..... | 49 |
| 6.2. Saran..... | 50 |
| DAFTAR PUSTAKA | |
| LAMPIRAN | |

DAFTAR TABEL

| Nomor | Judul | Halaman |
|-------|-------------------------------------------------------------------------------------|---------|
| 2.1 | Rangkuman Penelitian Terdahulu Metode Regresi <i>Ridge</i> | 7 |
| 3.1 | Analisis Ragam untuk Pengujian Parameter | 24 |
| 3.2 | Indikator Komponen IPM | 26 |
| 4.1 | Definisi Operasional Variabel | 27 |
| 5.1 | Hasil Pengujian Parameter Secara <i>Overall</i> | 38 |
| 5.2 | Hasil Pengujian Parameter Secara Parsial | 39 |
| 5.3 | Tabel Multikolinearitas | 40 |
| 5.4 | Data Transformasi | 41 |
| 5.5 | Nilai VIF $\hat{\beta}(c)$ dengan Berbagai Nilai c | 42 |
| 5.6 | Nilai Tetapan Bias | 44 |
| 5.7 | Koefisien Regresi Hasil MKT dan Regresi <i>Ridge</i> | 45 |
| 5.8 | Perbandingan Nilai Variansi, Bias, MSE, R^2 dan <i>Adjusted</i> R^2 | 45 |
| 5.9 | Uji <i>Overall</i> Regresi <i>Ridge</i> | 46 |
| 5.10 | Hasil Pengujian Parameter Secara Parsial Metode (Hoerl, Kennard & Baldwin, 1975) | 47 |
| 5.11 | Hasil Pengujian Parameter Secara Parsial Metode (Lawless & Wang, 1976) | 48 |

DAFTAR GAMBAR

| Nomor | Judul | Halaman |
|--------------|-------------------------------------------------|----------------|
| 1.1 | IPM Jawa Tengah Tahun 2010-2016 | 4 |
| 4.1 | Tahapan Penelitian | 30 |
| 5.1 | Deskriptif IPM Provinsi Jawa Tengah Tahun 2016 | 31 |
| 5.2 | Deskriptif PDRB Provinsi Jawa Tengah Tahun 2016 | 32 |
| 5.3 | Deskriptif AHH Provinsi Jawa Tengah Tahun 2016 | 33 |
| 5.4 | Deskriptif HLS Provinsi Jawa Tengah Tahun 2016 | 34 |
| 5.5 | Deskriptif RLS Provinsi Jawa Tengah Tahun 2016 | 35 |
| 5.6 | Plot Linier Variabel AHH, HLS dengan IPM | 36 |
| 5.7 | Plot Linier Variabel RLS, RMG dengan IPM | 36 |
| 5.8 | Plot Linier Variabel APS, PDRB dengan IPM | 37 |
| 5.9 | Plot Linier Variabel TPAK, KP dengan IPM | 37 |
| 5.10 | <i>Ridge Trace</i> | 44 |

DAFTAR LAMPIRAN

| Nomor | Judul | Halaman |
|--------------|---------------------------------------------|----------------|
| 1 | Data IPM di Provinsi Jawa Tengah Tahun 2016 | 53 |
| 2 | Data Transformasi | 54 |
| 3 | <i>Output</i> Program R | 55 |

DAFTAR ISTILAH

| | |
|------|-----------------------------------------|
| AHH | : Angka Harapan Hidup |
| AKB | : Angka Kematian Bayi |
| AMH | : Angka Melek Huruf |
| APS | : Angka Partisipasi Sekolah SMP/MTS |
| JSK | : Jumlah Sarana Kesehatan |
| HLS | : Harapan Lama Sekolah |
| IPM | : Indeks Pembangunan Manusia |
| KP | : Kepadatan Penduduk |
| PDRB | : Produk Domestik Regional Bruto |
| PK | : Penduduk Usia > 15 Tahun Yang Bekerja |
| PPK | : Pengeluaran Per Kapita |
| RLS | : Rata-Rata Lama Sekolah |
| RMG | : Rasio Murid Guru SMP/MTS |
| RMS | : Rasio Murid Sekolah SMP/MTS |
| RTAB | : Rumah Tangga Dengan Akses Air Bersih |
| TPAK | : Tingkat Partisipasi Angkatan Kerja |

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, Maret 2018


METERAI
TEMPEL
BBE45AFF009516887
6000
RUPIAH
amroni Rosyadi

**PENERAPAN METODE REGRESI *RIDGE* UNTUK MENGATASI
MASALAH MULTIKOLINEARITAS PADA KASUS INDEKS
PEMBANGUNAN MANUSIA DI PROVINSI JAWA TENGAH**

M Zamroni Rosyadi

Program Studi Statistika Fakultas MIPA

Universitas Islam Indonesia

INTISARI

Regresi linier berganda adalah suatu teknik dalam metode statistika yang digunakan untuk melihat pengaruh dua atau lebih variabel prediktor terhadap variabel respon. Salah satu asumsi pada analisis regresi linier berganda adalah tidak adanya multikolinearitas atau tidak adanya korelasi antara variabel-variabel prediktor di dalam model regresi. Ada beberapa metode yang digunakan untuk mengatasi masalah multikolinearitas diantaranya yaitu regresi *ridge*. Metode regresi *ridge* diperoleh dengan cara yang sama seperti metode kuadrat terkecil, yaitu dengan meminimumkan jumlah kuadrat sisaan. Regresi *ridge* menambahkan tetapan bias pada kuadrat terkecil sehingga koefisien berkurang dan mendekati nol. Terdapat beberapa metode yang digunakan untuk menentukan nilai tetapan bias, diantaranya yaitu menurut (Hoerl, Kennard & Baldwin, 1975) dan (Lawless & Wang, 1976). Hasil penelitian diperoleh metode terbaik untuk menentukan nilai tetapan bias dalam mengatasi masalah multikolinearitas pada kasus indeks pembangunan manusia di Provinsi Jawa Tengah tahun 2016 yaitu menggunakan metode regresi *ridge* (Lawless & Wang, 1976) karena memiliki nilai bias dan MSE yang lebih kecil serta nilai *adjusted R*² yang lebih besar dibandingkan metode regresi *ridge* (Hoerl, Kennard & Baldwin, 1975) sehingga didapatkan model persamaan regresi *ridge* (Lawless & Wang, 1976) yaitu: $\hat{Y} = -319.284 + 2.65431686 X_1 + 5.978742361 X_2 + 7.4803488 X_3 + 0.005698536 X_6$.

Kata Kunci : *Regresi Linier Berganda, Metode Kuadrat Terkecil, Regresi Ridge, Tetapan Bias*

**IMPLEMENTATION OF RIDGE REGRESSION METHOD TO SOLVE
MULTICOLLINEARITY PROBLEMS IN CASE OF HUMAN DEVELOPMENT
INDEX IN CENTRAL JAVA PROVINCE**

By: M Zamroni Rosyadi

Departement of Statistics Faculty of Mathematics and Natural Sciences

Islamic University of Indonesia

ABSTRACT

Multiple linear regression is a technique in statistical methods used to see the effect of two or more predictor variables on response variables. One of the assumptions in multiple linear regression analysis is the absence of multicollinearity or the absence of correlation between predictor variables in the regression model. There are several methods used to overcome multicollinearity problems such as ridge regression. The ridge regression method is obtained in the same way as the least squares method, that is, by minimizing the sum of the squared squares. The ridge regression adds the bias constant to the least squares so that the coefficient is reduced and close to zero. There are several methods used to determine the value of bias constant according to (Hoerl, Kennard & Baldwin) and according to (Lawless & Wang). The result of the research is based on the best method to determine the values of bias constant in the problem of multicollinearity in human development case index in Central Java Province in 2016 using the method of determining the value of bias constant according to Lawless & Wang because it has smaller bias and MSE adjusted R² is greater than (Hoerl, Kennard & Baldwin), the model found is the regression of ridge according to (Lawless & Wang, 1976) is: $\hat{Y} = -319.284 + 2.65431686 X_1 + 5.978742361 X_2 + 7.4803488 X_3 + 0.005698536 X_6$.

Keywords: *Multiple Linear Regression, Least Square Method, Ridge Regression, Bias Constant*

BAB I

PENDAHULUAN

1.1. Latar Belakang

Analisis data diartikan sebagai upaya untuk mengolah data menjadi informasi, sehingga karakteristik atau sifat-sifat data tersebut dapat dengan mudah dipahami dan bermanfaat untuk menjawab masalah-masalah yang berkaitan dengan kegiatan penelitian. Ada beberapa teknik statistika yang dapat digunakan untuk menganalisis data. Salah satu metode analisis data yang seringkali digunakan adalah analisis regresi. Analisis regresi merupakan suatu metode analisis statistika yang digunakan untuk melihat pengaruh antara satu atau lebih variabel prediktor terhadap variabel respon, dimana variabel prediktor adalah variabel yang mempengaruhi dan variabel respon adalah variabel yang dipengaruhi (Ghozali, 2013). Terdapat dua jenis model regresi linear yaitu model regresi linear sederhana dan berganda. Model regresi linear sederhana digunakan jika peneliti hendak mengetahui hubungan atau pengaruh satu variabel prediktor terhadap variabel respon. Jika seorang peneliti hendak mengkaji hubungan atau pengaruh dua atau lebih variabel prediktor terhadap variabel respon, maka model regresi yang digunakan adalah model regresi linear berganda (*multiple linear regression model*). Model regresi linear sederhana maupun model regresi linear berganda dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya menggunakan metode kuadrat terkecil (*ordinary least square*). Dalam statistika sebuah model regresi dikatakan cocok atau baik jika garis regresi yang dihasilkan untuk melakukan estimasi atau peramalan dari sebaran data menghasilkan error yang terkecil. Untuk melakukan analisis regresi harus dipenuhi berbagai asumsi klasik, antara lain data tidak mengalami autokorelasi, heteroskedastisitas dan multikolinearitas.

Permasalahan yang sering muncul adalah multikolinearitas yaitu terjadinya korelasi yang cukup tinggi antara variabel-variabel prediktor (Astuti, 2014). Multikolinearitas mengakibatkan determinan matriks $X'X$ mendekati nol sehingga menyebabkan matriks hampir singular yang berakibat nilai penduga parameter menjadi tidak stabil. Akibat lain yang ditimbulkan adalah nilai penduga parameter bisa mempunyai tanda yang salah atau jauh lebih besar daripada yang diperkirakan (Draper & Smith, 1992).

Multikolinearitas dalam model regresi linear dapat dideteksi dengan beberapa cara, diantaranya dengan menghitung nilai *Variance Inflation Factor* (VIF) dan *Tolerance* (TOL). Jika terdapat pelanggaran asumsi multikolinearitas, terdapat beberapa prosedur yang dapat digunakan untuk mengatasinya, seperti menambahkan data, menghilangkan satu atau beberapa variabel prediktor yang memiliki korelasi tinggi dari model regresi dan menggunakan metode analisis yang lain seperti regresi *ridge* (Ghozali, 2013).

Regresi *ridge* diajukan sebagai suatu cara untuk mengatasi penyimpangan multikolinearitas. Keuntungan penggunaan regresi *ridge* dibandingkan metode lain yaitu regresi *ridge* mengurangi dampak multikolinearitas dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga regresi linear berganda (Pratiwi, 2016) Metode regresi *ridge* diperoleh dengan cara yang sama seperti metode kuadrat terkecil, yaitu dengan meminimumkan jumlah kuadrat sisaan. Regresi *ridge* menambahkan kendala (tetapan bias) pada kuadrat terkecil sehingga koefisien berkurang dan mendekati nol (Hastie, Robert & Jerome 2008). Ada beberapa cara yang dapat digunakan untuk menentukan nilai tetapan bias yaitu:

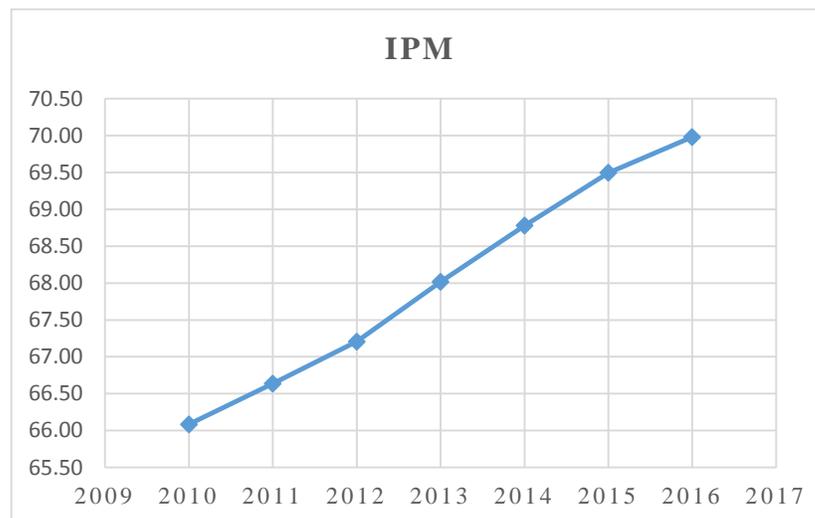
1. (Hoerl & Kennard, 1970).
2. (Hoerl, Kennard & Baldwin, 1975).
3. (Lawless & Wang, 1976).
4. (McDonaled & Galarnau, 1975).
5. (Dempster, Schatzoff & Wermuth, 1977).

Diantara 5 metode tersebut, (Hoerl, Kennard & Baldwin, 1975) dan (Lawless & Wang, 1976) merupakan metode yang paling sering digunakan untuk menentukan nilai tetapan bias (Astuti, 2014). Berdasarkan latar belakang tersebut, masalah multikolinearitas yang terjadi pada kasus IPM di Provinsi Jawa Tengah akan diselesaikan dengan menggunakan metode regresi *ridge*. IPM didefinisikan sebagai proses perluasan pilihan bagi penduduk (*enlarging people choice*). IPM merupakan indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia (masyarakat/penduduk). IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya. IPM diperkenalkan oleh UNDP pada tahun 1990 dan metode penghitungan direvisi pada tahun 2010. BPS mengadopsi perubahan metodologi penghitungan IPM yang baru pada tahun 2014 dan melakukan *backcasting* sejak tahun 2010.

IPM dibentuk oleh tiga dimensi dasar, yaitu umur panjang dan hidup sehat (*a long and healthy life*), pengetahuan (*knowledge*) dan standar hidup layak (*decend standard of living*). Umur panjang dan hidup sehat digambarkan oleh AHH yaitu jumlah tahun yang diharapkan dapat dicapai oleh bayi yang baru lahir untuk hidup dengan asumsi bahwa pola angka kematian menurut umur pada saat kelahiran sama sepanjang usia bayi. Pengetahuan diukur melalui indikator RLS dan HLS. RLS adalah rata-rata lamanya (tahun) penduduk usia 25 tahun ke atas dalam menjalani pendidikan formal. HLS didefinisikan sebagai lamanya (tahun) sekolah formal yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang. Standar hidup layak digambarkan oleh pengeluaran per kapita disesuaikan, yang ditentukan dari nilai pengeluaran per kapita dan paritas daya beli.

IPM dihitung berdasarkan rata-rata geometrik indeks kesehatan, indeks pengetahuan dan indeks pengeluaran. Penghitungan ketiga indeks ini dilakukan dengan melakukan standarisasi dengan nilai minimum dan maksimum masing-masing komponen indeks.

IPM merupakan indikator yang digunakan untuk melihat perkembangan pembangunan dalam jangka panjang. Untuk melihat kemajuan pembangunan manusia, terdapat dua aspek yang perlu diperhatikan, yaitu kecepatan dan status pencapaian. Secara umum, IPM Jawa Tengah terus mengalami kemajuan selama periode 2010 hingga 2016. IPM Jawa Tengah meningkat dari 66,08 pada tahun 2010 menjadi 69,98 pada tahun 2016. Selama periode tersebut, IPM Jawa Tengah rata-rata tumbuh sebesar 1,03 persen per tahun. Pada periode 2014-2015, IPM Jawa Tengah meningkat 0,71 poin. Peningkatan pada periode tersebut lebih rendah apabila dibandingkan dengan periode 2013-2014 yang naik sebesar 0,76 poin. Meskipun selama periode 2010 hingga 2016 IPM Jawa Tengah menunjukkan kemajuan yang besar, status IPM Jawa Tengah masih stagnan. Hingga saat ini, IPM Jawa Tengah masih berstatus “sedang” dan masih sama sejak tahun 2010.



Gambar 1.1 IPM Jawa Tengah Tahun 2010-2016 (Data diolah)

Dalam Penelitian Ayunanda & Ismaini (2013) dengan menggunakan regresi data panel terdapat beberapa variabel prediktor yang mempengaruhi IPM yaitu: RMG, RMS, APS, JSK, RTAB, KP, TPAK dan PDRB. Sedangkan penelitian yang dilakukan oleh Kartika Ayu, Maria & Rahma (2013) yang menggunakan metode *Partial Least Square Regression* (PLS-R) dalam regresi logistik ordinal menyimpulkan bahwa

variabel yang mempengaruhi IPM adalah AHH, AMH, RLS, PPK, AKB dan PK. Berdasarkan uraian tersebut maka variabel-variabel prediktor AHH, HLS, RLS, RMG, APS, PDRB, TPAK dan KP.

1.2. Rumusan Masalah

Dari uraian latar belakang yang telah dikemukakan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana hasil perbandingan metode penentuan nilai tetapan bias menurut (Hoerl, Kennard & Baldwin, 1975) dan menurut (Lawless & Wang, 1976)?
2. Bagaimana contoh penerapan regresi *ridge* untuk mengatasi multikolinearitas pada kasus IPM di Provinsi Jawa Tengah tahun 2016?

1.3. Jenis Penelitian dan Metode Analisis

Jenis penelitian dalam skripsi ini adalah penelitian aplikasi yang mengacu pada topik yang berjudul “Penggunaan Metode *Ridge Trace* dan *Variance Inflation Factors* (VIF) pada Regresi *Ridge*” yang dilakukan oleh Agriska Prenadita (2011). Pada topik tersebut dibahas tentang penyelesaian masalah multikolinearitas menggunakan metode regresi *ridge* pada kasus jumlah permintaan ayam di AS tahun 1982-2014. Berdasarkan uraian tersebut, maka peneliti akan menerapkan data kasus IPM di Provinsi Jawa Tengah tahun 2016 yang mengandung masalah multikolinearitas untuk selanjutnya diselesaikan dengan menggunakan metode regresi *ridge*.

1.4. Tujuan Penelitian

Sesuai dengan latar belakang masalah dan perumusan masalah yang telah diuraikan diatas, maka penelitian ini mempunyai tujuan sebagai berikut:

1. Menjelaskan hasil perbandingan metode penentuan nilai tetapan bias menurut (Hoerl, Kennard & Baldwin, 1975) dan menurut (Lawless & Wang, 1976).
2. Menunjukkan contoh penerapan regresi *ridge* untuk mengatasi multikolinearitas pada kasus IPM di Provinsi Jawa Tengah tahun 2016.

1.5. Batasan Masalah

Dalam penulisan skripsi ini, batasan masalah yang digunakan oleh penulis yaitu data di asumsikan telah memenuhi asumsi bahwa residual data berdistribusi normal, nilai variansi dari residual memiliki variansi yang konstan (homoskedastisitas) dan residual tidak saling berkorelasi (non-autokorelasi).

1.6. Manfaat Penelitian

Adapun manfaat yang diharapkan pada penelitian ini adalah memberikan pengetahuan tentang regresi *ridge* sebagai salah satu metode estimasi alternatif untuk mengatasi masalah multikolinearitas yang terjadi pada regresi linear berganda.

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Sebelumnya

Penelitian terdahulu sangat penting bagi penulis untuk mengetahui hubungan antara penelitian yang dilakukan sebelumnya dengan penelitian yang dilakukan saat ini dan untuk menghindari duplikasi. Penulisan penelitian terdahulu untuk menunjukkan bahwa penelitian yang dilakukan tersebut mempunyai manfaat sehingga dapat diketahui kontribusi penelitian tersebut terhadap ilmu pengetahuan saat ini. Penulisan penelitian ini berangkat dari beberapa penelitian dengan tema serupa yang telah dilakukan.

Penelitian tersebut salah satunya adalah Dereny dan Rashwan (2011) dalam jurnalnya yang berjudul “*Solving Multicollinearity Problem Using Ridge Regression Models*” menjelaskan beberapa jenis metode yang berbeda dari regresi *ridge*, antara lain dengan *Generalized Ridge Regression (GRR)*, *Ordinary Ridge Regression (ORR)* dan *Directed Ridge Regression (DRR)*. Hasil dari penelitian yang dilakukan menunjukkan bahwa estimator dari regresi *ridge* lebih baik dibanding estimator OLS saat ditemukannya kasus multikolinearitas.

Kemudian Prenadita (2011) dengan judul “Penggunaan Metode *Ridge Trace* dan *Variance Inflation Factors (VIF)* pada Regresi *Ridge*” untuk mengatasi multikolinearitas menggunakan metode regresi *ridge*. Hasil yang didapat yaitu dengan menggunakan *ridge trace*, yaitu dengan menambah tetapan bias c pada diagonal matriks $X'X$ maka akan diperoleh pada nilai c tertentu. Akibatnya, nilai *Variance Inflation Factors (VIF)* akan relatif dekat dengan 1 (kurang dari 10) sehingga koefisien β lebih stabil.

Selanjutnya yaitu penelitian yang berjudul “*Partial Least Square (PLS)* dan *Principal Component Regression (PCR)* untuk Regresi Linear dengan

Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Kabupaten Gunung Kidul” yang dilakukan Astuti (2014) dengan tujuan untuk mendapatkan hasil perbandingan metode *Partial Least Square (PLS)* dan *Principal Component Regression (PCR)* dalam penanggulangan masalah multikolinearitas. Dari hasil yang diperoleh bahwa metode *Partial Least Square (PLS)* memberikan hasil yang lebih baik dibandingkan metode *Principal Component Regression (PCR)*.

Puri (2014) dalam skripsinya yang berjudul “Aplikasi *Generalized Ridge Regression* untuk Menangani Masalah Multikolinearitas” membahas mengenai pengembangan terhadap regresi *ridge* dengan modifikasi, yaitu penambahan tetapan bias c yang berbeda ke dalam matriks korelasi $X'X$.

Berdasarkan penelitian-penelitian yang telah dilakukan, maka peneliti dengan tugas akhir yang berjudul “Penerapan Metode Regresi *Ridge* untuk Menangani Masalah Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Provinsi Jawa Tengah” membahas mengenai perbandingan metode dalam memilih nilai tetapan bias dengan menggunakan dua metode, yaitu menurut (Hoerl, Kennard & Baldwin, 1975) dan (Lawless & Wang, 1976). Hasil yang diperoleh yaitu metode pemilihan nilai tetapan bias menurut (Lawless & Wang, 1976) merupakan metode terbaik dalam mengatasi permasalahan multikolinearitas pada kasus IPM di Provinsi Jawa Tengah tahun 2016 karena memiliki nilai bias, MSE yang lebih kecil dan nilai *adjusted R²* yang lebih besar dibandingkan metode (Hoerl, Kennard & Baldwin, 1975). Rangkuman penelitian terdahulu tersebut dapat disajikan dalam **Tabel 2.1**.

Tabel 2.1 Rangkuman Penelitian Terdahulu Metode Regresi *Ridge*

| No | Nama/Tahun | Metode | Judul Penelitian | Hasil |
|----|---------------------------|----------------------|------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| 1 | Dereny dan Rashwan (2011) | Regresi <i>ridge</i> | <i>Solving Multicollinearity Problem Using Ridge Regression Models</i> | Estimator dari regresi <i>ridge</i> lebih baik dibanding estimator OLS saat ditemukannya kasus multikolinearitas |

| No | Nama/Tahun | Metode | Judul Penelitian | Hasil |
|----|------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2 | Prenadita (2011) | <i>Ridge Trace</i> dan <i>Variance Inflation</i> | Penggunaan Metode <i>Ridge Trace</i> dan <i>Variance Inflation Factors (VIF)</i> pada Regresi Ridge“ untuk mengatasi multikolinearitas menggunakan metode regresi <i>ridge</i> | Dengan menggunakan <i>ridge trace</i> , yaitu dengan menambah tetapan bias c pada diagonal matriks $X'X$ maka akan diperoleh pada nilai c tertentu. Akibatnya, nilai <i>Variance Inflation Factors (VIF)</i> akan relatif dekat dengan 1 (kurang dari 10) sehingga koefisien β lebih stabil. |
| 3 | Astuti (2014) | <i>Partial Least Square (PLS)</i> dan <i>Principal Component Regression (PCR)</i> | <i>Partial Least Square (PLS)</i> dan <i>Principal Component Regression (PCR)</i> untuk Regresi Linear dengan Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Kabupaten Gunung Kidul | Metode <i>Partial Least Square (PLS)</i> memberikan hasil yang lebih baik dibandingkan metode <i>Principal Component Regression (PCR)</i> |
| 4 | Puri (2014) | Regresi <i>ridge</i> | Aplikasi <i>Generalized Ridge Regression</i> untuk Menangani Masalah Multikolinearitas | Hasil estimasi menggunakan analisis <i>Generalized Ridge Regression</i> lebih baik dibandingkan dengan analisis regresi biasa saat terdeteksi ada multikolinearitas |

BAB III

LANDASAN TEORI

3.1. Model Regresi Linear Berganda

Menurut Sembiring (2003), secara umum model regresi linear berganda dengan variabel respon (Y) yang merupakan fungsi linear dari k variabel prediktor (X_1, X_2, \dots, X_k), dapat ditulis:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (3.1)$$

dimana:

Y = variabel respon

X_k = variabel prediktor, ($k = 1, 2, \dots, p$)

β_0 = *intercept*

β_k = koefisien regresi pada variabel X_k , ($k = 1, 2, \dots, p$)

ε = variabel pengganggu/residual

Jika dipunyai sampel random berukuran n , yaitu $(Y_i, X_{1i}, X_{2i}, \dots, X_{ki})$ di mana $i = 1, 2, \dots, n$ dari sebuah populasi, maka model regresinya dapat ditulis :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (3.2)$$

Sepertinya halnya analisis regresi linear sederhana, untuk mendapatkan model yang baik, yang mana model dapat menggambarkan keadaan yang sebenarnya dalam analisis regresi berganda perlu diperhatikan lima asumsi regresi linear. Karena analisis regresi melibatkan lebih dari satu variabel prediktor, maka perlu asumsi tambahan berkaitan dengan hubungan antar variabel prediktor, yaitu salah satunya adalah tidak adanya hubungan linear di antara variabel-variabel prediktor dalam model regresi.

3.2. Metode Kuadrat Terkecil (*Ordinary Least Square*)

Metode kuadrat terkecil (OLS) adalah salah satu metode yang dapat digunakan untuk mengestimasi parameter. OLS merupakan metode estimasi fungsi regresi yang paling sering digunakan. Kriteria OLS adalah "*Line of Best Fit*" atau dengan kata lain

jumlah kuadrat dari deviasi antara titik-titik observasi dengan garis regresi adalah minimum. Analisis regresi dengan metode OLS akan memberikan hasil yang *Best Linear Unbiased Estimator* (BLUE) jika memenuhi semua asumsi klasik dengan $mean E\{\varepsilon_i\} = 0$ dan variansi $\sigma^2\{\varepsilon_i\} = \sigma^2$, ε_i dan ε_j tidak berkorelasi sehingga kovariansinya $E\{\varepsilon_i, \varepsilon_j\} = 0$, untuk semua nilai i dan j , $i \neq j$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$. (Quadratullah, 2013).

Model regresi linear berganda diatas dapat dinyatakan dalam bentuk matriks berikut:

$$Y = X\beta + \varepsilon \quad (3.3)$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan Y adalah vektor $n \times 1$ dari observasi-observasi pada variabel respon, X adalah matriks $n \times (k+1)$ dari observasi-observasi pada k -variabel prediktor, β adalah vektor $(k+1) \times 1$ dari koefisien regresi dan ε adalah vektor $i \times 1$ dari error dengan $\varepsilon_i \sim NID(0, \sigma^2)$.

Jika diasumsikan X dan Y telah dipusatkan dan diskalakan sehingga $1'X = 0$ dan $X'X = 1$ maka $X'X$ dan $X'Y$ adalah matriks korelasi dari koefisien-koefisien.

Estimator kuadrat terkecil untuk β adalah:

$$\hat{\beta} = (X'X)^{-1}X'Y \quad (3.4)$$

Bukti:

$$Y = X\beta + \varepsilon$$

$$\varepsilon = Y - X\beta$$

$$\begin{aligned} \varepsilon'\varepsilon &= (Y - X\beta)'(Y - X\beta) \\ &= Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \end{aligned}$$

Catatan: $\beta'X'Y$ adalah matriks 1×1 atau suatu skalar dan nilai transposenya $(\beta'X'Y)' = Y'X\beta$ adalah skalar yang sama.

Dalam metode kuadrat terkecil harus dipenuhi:

$$\left. \frac{\partial \varepsilon^t \varepsilon}{\partial \beta} \right|_{\hat{\beta}} = -2X^t Y + 2X^t X \hat{\beta} = 0$$

sehingga diperoleh :

$$X^t X \hat{\beta} = X^t Y$$

$$\hat{\beta} = (X^t X)^{-1} X^t Y$$

Matriks $X^t X$ dari persamaan (3.4) diatas adalah matriks simetris $(k+1) \times (k+1)$ yang mana elemen-elemen diagonalnya adalah jumlah kuadrat dari elemen-elemen kolom dalam matriks X dan elemen-elemen diluar diagonal adalah jumlahan hasil kali produk dari elemen-elemen dalam kolom yang sama. Pada dasarnya $X^t X$ mempunyai peranan penting dalam sifat-sifat estimator β dan sering menjadi faktor utama dalam kesuksesan atau kegagalan estimasi kuadrat terkecil.

3.3. Sifat-Sifat Estimator Kuadrat Terkecil

Metode kuadrat terkecil memiliki suatu estimator yang mempunyai sifat-sifat yang sangat baik. Sifat-sifat tersebut antara lain yaitu:

Sifat estimator $\hat{\beta} = (X^t X)^{-1} X^t Y$ (Bain, 1992).

1. Linear

$$\begin{aligned} \hat{\beta} &= (X^t X)^{-1} X^t Y \\ &= (X^t X)^{-1} X^t X \beta + \varepsilon \\ &= (X^t X)^{-1} X^t X \beta + (X^t X)^{-1} X^t \varepsilon \\ &= I \beta + (X^t X)^{-1} X^t \varepsilon \end{aligned}$$

2. Tak Bias

$$E(\hat{\beta}) = \beta \tag{3.5}$$

Bukti:

$$\begin{aligned} E(\hat{\beta}) &= E[(X^t X)^{-1} X^t Y] \\ &= E[(X^t X)^{-1} X^t (X \beta + \varepsilon)] \\ &= E[(X^t X)^{-1} X^t X \beta + (X^t X)^{-1} X^t \varepsilon] \end{aligned}$$

$$\begin{aligned}
&= \beta + E[(X'X)^{-1}X'\varepsilon] \\
&= \beta + (X'X)^{-1}X'E[\varepsilon] \\
&= \beta + 0 \\
&= \beta
\end{aligned}$$

jadi $\hat{\beta}$ merupakan penaksir tak bias dari β .

3. Variansi minimum

Cara menunjukkan bahwa semua β_i dalam vektor $\hat{\beta}$ adalah penaksir-penaksir terbaik (*best estimator*), harus dibuktikan bahwa $\hat{\beta}$ mempunyai variansi yang terkecil atau minimum diantara variansi estimator tak bias linear yang lain.

$$\begin{aligned}
\text{Var}(\hat{\beta}) &= E[(\hat{\beta} - \beta)^2] & (3.6) \\
&= E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] \\
&= E[\{(X'X)^{-1}X'\varepsilon\}\{(X'X)^{-1}X'\varepsilon\}'] \\
&= E[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}] \\
&= (X'X)^{-1}X'E[\varepsilon\varepsilon']X(X'X)^{-1} \\
&= (X'X)^{-1}X'\sigma^2IX(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\
&= \sigma^2(X'X)^{-1}
\end{aligned}$$

Akan ditunjukkan $\text{var}(\hat{\beta}) \leq \text{var}(\hat{\beta}^*)$.

Misal $\hat{\beta}^*$ adalah estimator linear yang lain dari β yang dapat ditulis sebagai

$$\hat{\beta}^* = [(X'X)^{-1}X' + c]Y \quad (3.7)$$

dengan c adalah matriks konstanta, sehingga

$$\begin{aligned}
\hat{\beta}^* &= [(X'X)^{-1}X' + c]Y \\
&= [(X'X)^{-1}X' + c](X\beta + \varepsilon) \\
&= (X'X)^{-1}X'X\beta + cX\beta + (X'X)^{-1}X'\varepsilon + c\varepsilon \\
&= I\beta + cX\beta + (X'X)^{-1}X'\varepsilon + c\varepsilon \\
&= \beta + cX\beta + (X'X)^{-1}X'\varepsilon + c\varepsilon
\end{aligned}$$

Karena diasumsikan $\hat{\beta}^*$ merupakan estimator tak bias dari β maka $E(\hat{\beta}^*)$ seharusnya β , dengan kata lain $cX\beta$ seharusnya merupakan matriks nol, atau $cX = 0$.

$$\text{Jadi diperoleh } \hat{\beta}^* - \beta = (X^tX)^{-1}X^t \varepsilon + c \varepsilon = ((X^tX)^{-1}X^t+c) \varepsilon \quad (3.8)$$

$$\begin{aligned} \text{Var}(\hat{\beta}^*) &= E[(\hat{\beta}^* - \beta)(\hat{\beta}^* - \beta)^t] \\ &= E[(X^tX)^{-1}X^t+c) \varepsilon \varepsilon^t(X(X^tX)^{-1}+c^t)] \\ &= ((X^tX)^{-1}X^t+c)E(\varepsilon \varepsilon^t)(X(X^tX)^{-1}+c^t) \\ &= \sigma^2((X^tX)^{-1}X^t+c)(X(X^tX)^{-1}+c^t) \\ &= \sigma^2((X^tX)^{-1}X^tX(X^tX)^{-1}+cX(X^tX)^{-1}+(X^tX)^{-1}X^tc^t+cc^t) \\ &= \sigma^2((X^tX)^{-1}+cc^t) \\ &= \text{var}(\hat{\beta}) + \sigma^2cc^t \end{aligned}$$

Persamaan di atas menunjukkan bahwa matriks variansi estimator linear tak bias $\hat{\beta}^*$ merupakan penjumlahan matriks variansi estimator OLS dengan σ^2cc^t . Secara matematis jadi terbukti bahwa $\text{var}(\hat{\beta}) \leq \text{var}(\hat{\beta}^*)$.

3.4. Ukuran Pemusatan dan Penskalaan (*Centering and Scaling*)

Pemusatan dan penskalaan data merupakan bagian dari membakukan (*standardized*) variabel. Modifikasi sederhana dari pembakuan atau standarisasi variabel ini adalah transformasi korelasi (*correlation transformation*). Pemusatan merupakan perbedaan antara masing-masing pengamatan dan rata-rata dari semua pengamatan untuk variabel. Sedangkan penskalaan meliputi gambaran pengamatan pada kesatuan (unit) standar deviasi dari pengamatan untuk variabel (Kutner, 2005).

Dalam hal ini yang akan dibakukan (distandarisasi) adalah model regresi linear berganda yang ditunjukkan pada model dibawah ini.

$$Y_i = \beta_0 + \beta_1X_{1i} + \beta_2X_{2i} + \dots + \beta_kX_{ki} + \varepsilon_i \quad (3.9)$$

Berikut ini merupakan pembakuan variabel terikat Y dan variabel bebas X_1, X_2, \dots, X_k .

$$\frac{Y_i - \bar{Y}}{S_Y}, \text{ dengan } S_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}} \quad (3.10)$$

$$\frac{X_{ij} - \bar{X}_j}{S_{Xj}}, \text{ dengan } S_{Xj} = \sqrt{\frac{\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}{n-1}} \quad j = 1, 2, \dots, k \quad (3.11)$$

Keterangan:

\bar{Y} = rata-rata dari Y

\bar{X}_j = rata-rata dari pengamatan X_j

S_Y = standar deviasi dari Y

S_{Xj} = standar deviasi dari X_j

Transformasi korelasi merupakan fungsi sederhana dari pembakuan variabel.

Sehingga melalui transformasi dapat diperoleh:

$$Y_i^* = \frac{Y_i - \bar{Y}_i}{\sqrt{n-1} S_Y} \quad (3.12)$$

$$X_{ij}^* = \frac{X_{ki} - \bar{X}_k}{\sqrt{n-1} S_{Xj}} \quad (3.13)$$

Berdasarkan transformasi peubah Y_i^* dan X_{ij}^* yang didefinisikan dengan transformasi korelasi pada model (3.12) dan (3.13) di atas diperoleh model regresi sebagai berikut

$$Y_i^* = \beta_1^* X_{1i} + \beta_2^* X_{2i} + \dots + \beta_k^* X_{ki} + \varepsilon_i \quad (3.14)$$

Model (3.14) di atas disebut sebagai model regresi yang baku (*standardized regression model*). Diantara parameter $\beta_1^*, \beta_2^*, \dots, \beta_k^*$ pada model regresi baku dengan parameter asli $\beta_1, \beta_2, \dots, \beta_k$ pada model regresi linear berganda yang biasa terdapat suatu hubungan linear. Hubungan antara kedua parameter dari dua model yang berbeda tersebut dijabarkan seperti di bawah ini (Kutner, 2005):

$$\beta_j = \left(\frac{S_Y}{S_{Xj}} \right) \beta_j^*, \quad j = 1, 2, \dots, k \quad (3.15)$$

$$\begin{aligned} \beta_0 &= \bar{Y} - \beta_1 \bar{X}_1 - \beta_2 \bar{X}_2 - \dots - \beta_k \bar{X}_k \\ &= \bar{Y} - \sum_{j=1}^k \beta_j \bar{X}_j \end{aligned} \quad (3.16)$$

Prosedur ini disebut dengan prosedur penskalaan.

Model (3.10) di atas dapat dibentuk menjadi :

$$Y_i = \beta_0 + \beta_1(X_{1i} - \bar{X}_1) + \beta_1 \bar{X}_1 + \beta_2(X_{2i} - \bar{X}_2) + \beta_2 \bar{X}_2 + \dots + \beta_k(X_{ki} - \bar{X}_k) + \beta_k \bar{X}_k + \varepsilon_i$$

$$= (\beta_0 + \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 + \dots + \beta_k \bar{X}_k) + \beta_1 (X_{1i} - \bar{X}_1) + \beta_2 (X_{2i} - \bar{X}_2) + \dots + \beta_k (X_{ki} - \bar{X}_k) + \varepsilon_i$$

Menurut rumus untuk mendapatkan β_0 yaitu:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}_1 - \beta_2 \bar{X}_2 - \dots - \beta_k \bar{X}_k$$

Maka berlaku:

$$\bar{Y} = \beta_0 + \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 + \dots + \beta_k \bar{X}_k$$

sehingga

$$\begin{aligned} Y_i - \bar{Y} &= (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i) - (\beta_0 + \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 + \dots + \beta_k \bar{X}_k) \\ &= \beta_1 (X_{1i} - \bar{X}_1) + \beta_2 (X_{2i} - \bar{X}_2) + \dots + \beta_k (X_{ki} - \bar{X}_k) + \varepsilon_i \end{aligned}$$

jika

$$y_i = Y_i - \bar{Y}$$

$$x_{1i} = X_{1i} - \bar{X}_1$$

$$x_{2i} = X_{2i} - \bar{X}_2$$

$$x_{ki} = X_{ki} - \bar{X}_k$$

Maka didapat model baru yaitu:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (3.17)$$

Prosedur untuk membentuk model pertama menjadi model terakhir disebut prosedur pemusatan. Prosedur ini mengakibatkan hilangnya β_0 (*intercept*) yang membuat perhitungan untuk mencari model regresi menjadi lebih sederhana. Keseluruhan dari prosedur di atas disebut prosedur pemusatan dan penskalaan.

3.5. Matriks Korelasi

Matriks korelasi dapat diperoleh dari hasil perkalian antara transpose matriks X dengan matriks X.

$$X^T X = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{n1} \\ X_{12} & X_{22} & \dots & X_{n2} \\ \vdots & \vdots & \dots & \vdots \\ X_{1n} & X_{2n} & \dots & X_{nn} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \dots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{bmatrix}$$

$$X'X = \begin{bmatrix} \sum X_{i1}^2 & \sum X_{i1} X_{i2} & \dots & \sum X_{i1} X_{in} \\ \sum X_{i1} X_{i2} & \sum X_{i2}^2 & \dots & \sum X_{i2} X_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{i1} X_{in} & \sum X_{i2} X_{in} & \dots & \sum X_{in}^2 \end{bmatrix}$$

Matriks $X'X$ yang telah di standarkan dapat ditulis sebagai berikut:

$$X'X = \begin{bmatrix} \sum X_{i1}^*{}^2 & \sum X_{i1}^* X_{i2}^* & \dots & \sum X_{i1}^* X_{in}^* \\ \sum X_{i1}^* X_{i2}^* & \sum X_{i2}^*{}^2 & \dots & \sum X_{i2}^* X_{in}^* \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{i1}^* X_{in}^* & \sum X_{i2}^* X_{in}^* & \dots & \sum X_{in}^*{}^2 \end{bmatrix}$$

dengan

$$\begin{aligned} \sum_{i=1}^n X_{i1}^*{}^2 &= \sum_{i=1}^n \left(\frac{X_{i1} - \bar{X}_1}{\sqrt{n-1}S_1} \right)^2 \\ &= \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2}{(n-1)S_1^2} \\ &= \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2}{(n-1) \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2}{(n-1)}} = 1 \\ \sum X_{i1}^* X_{i2}^* &= \sum_{i=1}^n \left(\frac{X_{i1} - \bar{X}_1}{\sqrt{n-1}S_1} \right) \left(\frac{X_{i2} - \bar{X}_2}{\sqrt{n-1}S_2} \right) \\ &= \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2)}{(n-1)S_1S_2} \\ &= \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2)}{(n-1) \sqrt{\frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2}{n-1}} \sqrt{\frac{\sum_{i=1}^n (X_{i2} - \bar{X}_2)^2}{n-1}}} \\ &= \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2)}{\sqrt{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2} \sqrt{\sum_{i=1}^n (X_{i2} - \bar{X}_2)^2}} \\ &= \frac{S_{12}}{\sqrt{S_{11}}\sqrt{S_{22}}} = r_{12} = r_{21} \end{aligned}$$

Sehingga matriks korelasi R adalah :

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p1} & \dots & 1 \end{bmatrix}; r_{12} = r_{21}, r_{13} = r_{31}, \dots, r_{1p} = r_{p1} \quad (3.18)$$

3.6. Multikolinearitas

Istilah multikolinearitas pertama kali ditemukan oleh Ragnar Frisch pada tahun 1934 yang berarti terdapat hubungan linear diantara beberapa variabel atau semua variabel prediktor dalam model regresi. Multikolinearitas adalah suatu kondisi dimana terjadi korelasi yang kuat diantara variabel-variabel prediktor (X) yang diikutsertakan dalam pembentukan model regresi linier. Multikolinearitas tidak mungkin terjadi apabila hanya terdapat satu variabel prediktor (X)

Dalam bentuk matriks, multikolinearitas adalah suatu kondisi buruk atau *ill condition* dari matriks $X'X$ yaitu suatu kondisi yang tidak memenuhi asumsi klasik. Jika multikolinearitas terjadi antara dua variabel atau lebih dalam suatu persamaan regresi, maka nilai perkiraan koefisien dari variabel yang bersangkutan menjadi tak berhingga, sehingga tidak mungkin lagi menduganya. Hal ini disebabkan $X'X$ menjadi singular atau $X'X$ mendekati nol. Dalam multikolinearitas terdapat dua jenis hubungan linear yang sempurna (multikolinearitas sempurna) dan hubungan linear kurang sempurna (multikolinearitas kurang sempurna).

3.6.1. Multikolinearitas Sempurna

Untuk hubungan yang terdiri dari k variabel, mencakup variabel prediktor X_1, X_2, \dots, X_k . Hubungan linier yang sempurna atau pasti terjadi jika berlaku hubungan berikut:

$$C_1X_1 + C_2X_2 + \dots + C_kX_k = 0 \quad (3.19)$$

dengan C_1, C_2, \dots, C_k merupakan bilangan konstan dan tidak seluruhnya nol atau paling tidak ada satu yang tidak sama dengan nol, yaitu $\exists C_j \neq 0$ ($j = 1, 2, \dots, k$).

3.6.2. Multikolinearitas Kurang Sempurna

Istilah multikolinearitas digunakan dalam arti lebih luas, yaitu mencakup hubungan linier sempurna dan juga dimana variabel-variabel prediktor X *interkorelasi*, akan tetapi tidak sempurna seperti hubungan berikut:

$$C_1X_2 + C_2X_2 + \dots + C_kX_k + \varepsilon_i = 0 \quad (3.20)$$

dengan ε_i = kesalahan pengguna (*standar error*).

Untuk mengetahui perbedaan antara multikolinearitas sempurna dan multikolinearitas kurang sempurna, diasumsikan $C_2 \neq 0$. Dapat ditunjukkan untuk setiap observasi ke- i persamaan (3.20) menjadi:

$$X_{2i} = -\frac{C_1}{C_2} X_{1i} - \frac{C_3}{C_2} X_{3i} - \dots - \frac{C_k}{C_2} X_{ki} \quad (3.21)$$

yang menunjukkan bagaimana variabel prediktor X_{2i} berhubungan linier secara sempurna dengan variabel lainnya secara keseluruhan atau bagaimana hubungan tersebut dapat diturunkan dari suatu hubungan linier antara variabel prediktor-variabel prediktor lainnya.

Diasumsikan $C_2 \neq 0$ maka persamaan (3.21) menjadi.

$$X_{2i} = -\frac{C_1}{C_2} X_{1i} - \frac{C_3}{C_2} X_{3i} - \dots - \frac{C_k}{C_2} X_{ki} - \frac{1}{C_2} \varepsilon_i \quad (3.22)$$

Persamaan diatas menunjukkan bahwa X_{2i} tidak berhubungan linier sempurna dengan variabel lainnya, sebab masih tergantung pada kesalahan pengganggu (ε_i) (Sumodiningrat, 1994).

3.7. Regresi Ridge

Prosedur jejak *ridge* (*ridge trace procedure*), pertama kali dikemukakan oleh A.E. Hoerl pada 1962, dibicarakan panjang lebar oleh A.E. Hoerl dan R.W. Kennard dalam “*Ridge Regression Biased Estimation for Nonorthogonal Problem*”, *Technometric*, 12, 1970, 55-67, makalah kedua dengan dilengkapi ilustrasi, muncul di halaman 69-82 pada terbitan yang sama. Prosedur ini ditujukan untuk mengatasi “kondisi buruk (*ill conditioned*)” yang diakibatkan oleh korelasi yang tinggi antara beberapa variabel bebas di dalam model, sehingga menyebabkan matrik $X'X$ -nya hampir singular, yang pada gilirannya menghasilkan nilai dugaan parameter model yang tidak stabil. (Misalnya, nilai dugaannya bisa mempunyai tanda yang salah atau jauh lebih besar daripada yang diperkirakan menurut pertimbangan fisik maupun praktis) (Draper & Smith, 1992).

Metode regresi *ridge* digunakan untuk mengurangi konsekuensi (dampak) dari multikolinearitas dengan cara menambahkan nilai (c) yang bias tetapi cenderung mempunyai rata-rata kuadrat *residual* yang lebih kecil daripada estimator yang diperoleh dengan metode OLS, sehingga diperoleh estimator regresi *ridge* yaitu:

$$\beta_R = (X'X + cI)^{-1} X'Y \quad (3.23)$$

Adapun sifat-sifat dari regresi *ridge* antara lain:

1. Nilai Ekspektasi dari Estimator Regresi *Ridge*

$$\begin{aligned} E(\beta_R) &= E[(X'X + cI)^{-1} X'Y] \\ &= E[(X'X + cI)^{-1} (X'X)(X'X)^{-1} X'Y] \\ &= E[(X'X + cI)^{-1} (X'X) \hat{\beta}] \\ &= E[(X'X + cI)^{-1} (X'X)] E(\hat{\beta}) \\ &= (X'X + cI)^{-1} X'X \beta \end{aligned}$$

Bias sebesar $(X'X + cI)^{-1} X'X$ dari β

Jika memanfaatkan hubungan antara $\hat{\beta}_R$ dengan β , maka diperoleh nilai ekspektasi dari estimator regresi *ridge* adalah sebagai berikut:

$$\begin{aligned} E(\hat{\beta}_R) &= [I - c(X'X + cI)^{-1}] \beta \\ &= \beta - c(X'X + cI)^{-1} \beta \end{aligned}$$

Bias sebesar $c(X'X + cI)^{-1} \beta$, $0 \leq c \leq \infty$ (3.24)

2. Menurut Montgomery dan Peck (1991) varian dari β_R dapat dinyatakan dalam bentuk matriks sebagai berikut.

$$V(\beta_R) = \sigma^2 (X'X + cI)^{-1} (X'X) (X'X + cI)^{-1} \quad (3.25)$$

3. Jumlah Kuadrat Kesalahan

$$\begin{aligned} SSE(\hat{\beta}_R) &= (Y - X\hat{\beta}_R)'(Y - X\hat{\beta}_R) \\ &= (Y - X\hat{\beta}_R)'(Y - X\hat{\beta}_R) + (\hat{\beta}_R - \hat{\beta})' X'X (\hat{\beta}_R - \hat{\beta}) \end{aligned} \quad (3.26)$$

Bukti :

$$\begin{aligned} SSE(\hat{\beta}_R) &= (Y - X\hat{\beta}_R)'(Y - X\hat{\beta}_R) \\ &= Y'Y - \hat{\beta}_R' X'Y - Y'X\hat{\beta}_R + \hat{\beta}_R' X'X\hat{\beta}_R \\ &= Y'Y - \hat{\beta}_R' (X'X) (X'X)' X'Y - Y'X\hat{\beta}_R (X'X) (X'X)' + \hat{\beta}_R' X'X\hat{\beta}_R \end{aligned}$$

4. Menurut Montgomery dan Peck (1991) rata-rata jumlah kuadrat dari regresi *ridge* adalah sebagai berikut.

$$\begin{aligned}
 \text{MSE}(\beta_R) &= \text{Var}(\beta_R) + (\text{bias}(\beta_R))^2 & (3.27) \\
 &= \sigma^2 \text{trace}[(X'X+cI)^{-1}(X'X)(X'X+cI)^{-1}] + [(-\beta c X'X+cI^{-1})^2] \\
 &= \sigma^2 \text{trace}[(X'X+cI)^{-1}(X'X)(X'X+cI)^{-1}] + c^2 \beta^t X'X+cI^{-1}]^{-2} \beta \\
 &= \sigma^2 \sum_{j=1}^p \frac{j}{\lambda_j+c} + c^2 \beta^t (X'X+cI)^{-2} \beta
 \end{aligned}$$

Masalah yang dihadapi dalam regresi *ridge* adalah penentuan nilai tetapan bias c . Beberapa cara untuk menghitung nilai c yaitu dengan menggunakan metode (Hoerl, Kennard & Baldwin) dan metode (Lawless & Wang). Untuk metode (Hoerl, Kennard & Baldwin, 1975), penentuan nilai tetapan bias dihitung dengan menggunakan rumus.

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}^t \beta} \quad (3.28)$$

Keterangan:

p = banyaknya parameter di luar β_0

$\hat{\sigma}^2$ = *mean square error* yang diperoleh dari metode OLS

$\hat{\beta}$ = vektor estimasi yang diperoleh dengan metode OLS

Selanjutnya untuk metode (Lawless & Wang, 1976), penentuan nilai tetapan bias dihitung dengan menggunakan rumus.

$$\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\hat{\beta}^t (X'X) \hat{\beta}} \quad (3.29)$$

Keterangan:

p = banyaknya parameter di luar β_0

$\hat{\sigma}^2$ = *mean square error* yang diperoleh dari metode OLS

$\hat{\beta}$ = vektor estimasi yang diperoleh dengan metode OLS

3.8. Mendeteksi Multikolinearitas dengan Metode *Ridge*

Ada dua metode untuk mendeteksi multikolinearitas yang berhubungan dengan regresi *ridge*. Metode pertama dihubungkan dengan efek multikolinearitas terhadap

error antara penduga kuadrat terkecil dan nilai sebenarnya dari koefisien regresi. Metode kedua berhubungan dengan ketidakstabilan penduga kuadrat terkecil dalam menghadapi perubahan kecil dalam data.

3.8.1. *Variance Inflation Factors (VIF)*

Variance Inflation Factors (VIF) merupakan salah satu indikator untuk mengukur besarnya multikolinearitas. VIF menunjukkan peningkatan ragam dari koefisien regresi yang disebabkan karena adanya ketergantungan linier peubah prediktor tersebut dengan peubah prediktor yang lain. Menurut Montgomery dan Peck (1991) *Variance Inflation Factors (VIF)* dapat dihitung dengan rumus sebagai berikut:

$$VIF = \frac{1}{1-R_j^2} \quad (3.30)$$

Dengan R_j^2 adalah koefisien determinasi ke- j , $j = 1, 2, \dots, k$. Jika nilai VIF lebih dari 5 atau 10 mengindikasikan adanya multikolinearitas (Novi Becti, 2016).

Cara mengetahui adanya multikolinearitas dengan melihat nilai VIF. Nilai VIF yang semakin besar akan menunjukkan multikolinearitas yang lebih kompleks. Jika nilai $VIF > 10$, maka dapat disimpulkan bahwa terdapat multikolinearitas (Soemartini, 2008).

3.8.2. *Ridge Trace*

Metode yang kedua untuk mendeteksi multikolinearitas dengan menggunakan *ridge trace* atau jejak *ridge*. Salah satu kesulitan utama dalam menggunakan regresi *ridge* adalah dalam menentukan nilai tetapan bias yang tepat (Hoerl & Kennard, 1970), pencipta regresi *ridge* menganjurkan untuk menggunakan suatu grafik yang mereka sebut sebagai *ridge trace*. *Ridge trace* adalah plot dari estimator regresi *ridge* dengan berbagai kemungkinan nilai tetapan bias c , konstanta c mencerminkan jumlah bias dalam estimator $\hat{\beta}(c)$. Bila $c = 0$ maka estimator $\hat{\beta}(c)$ akan bernilai sama dengan kuadrat terkecil β , tetapi cenderung lebih stabil daripada estimator kuadrat terkecil.

Ridge trace menggambarkan koefisien regresi *ridge* sebagai fungsi dari c . Pemilihan tetapan bias c merupakan masalah yang perlu diperhatikan. Tetapan bias yang diinginkan adalah tetapan bias yang menghasilkan bias relatif kecil dan menghasilkan koefisien yang relatif stabil.

3.9. Pengujian Signifikan Parameter Model

Apabila model telah ditentukan, maka dilakukan pengujian signifikan parameter model regresi. Pengujian ini bertujuan untuk mengetahui apakah parameter yang terdapat dalam model regresi telah menunjukkan hubungan yang tepat antara variabel-variabel prediktor dengan variabel respon. Pengujian signifikan parameter model regresi juga sebagai salah satu cara mengevaluasi seberapa baik model regresi yang diperoleh. Ada dua tahapan dalam pengujian signifikansi parameter model regresi, yaitu pengujian secara serentak atau *overall* dan pengujian secara parsial.

Pengujian parameter secara *overall* digunakan untuk mengevaluasi pengaruh semua variabel prediktor terhadap variabel respon. Hipotesis untuk pengujian *overall* yang digunakan adalah :

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_j = 0, \text{ (Model regresi tidak signifikan)}$$

$$H_1 : \exists \beta_j \neq 0, j = 0, 1, \dots, k \text{ (Model regresi signifikan)}$$

Statistik uji yang digunakan dalam pengujian parameter secara *overall* adalah statistik F_{hitung} dengan.

$$F_{hitung} = KTR/KTS \quad (3.31)$$

dengan

KTR : Kuadrat Tengah Regresi

KTS : Kuadrat Tengah Sisaan

Tabel analisis ragam dapat disusun sebagai berikut (Qudratullah, 2013) :

Tabel 3.1 Analisis Ragam untuk Pengujian Parameter

| Sumber | Derajat Bebas (db) | Jumlah Kuadrat (JK) | Kuadrat Tengah (KT) | F_{hitung} |
|---------|--------------------|----------------------------------------------|---------------------------|--------------------------------|
| Regresi | k | $JKR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ | $KTR = \frac{JKR}{k}$ | $F_{hitung} = \frac{KTR}{KTS}$ |
| Sisaan | $n-k-1$ | $JKS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ | $KTS = \frac{JKS}{n-k-1}$ | |
| Total | $n-1$ | $JKT = \sum_{i=1}^n (y_i - \bar{y})^2$ | | |

Statistik uji yang digunakan dalam pengujian parameter secara *overall* adalah statistik F_{hitung} dengan daerah kritis yaitu:

$$H_0 \text{ ditolak jika } F_{hitung} > F_{tabel}$$

dengan $F_{tabel} = F_{(k, n-k-1, \alpha)}$

Untuk pengujian parameter secara parsial menggunakan statistik t_{hitung} dengan.

$$t_{hitung} = \frac{\hat{\beta}}{s(\hat{\beta})} \quad (3.32)$$

$$\text{dimana } s(\hat{\beta}) = \sqrt{\frac{s^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (3.33)$$

Pengujian ini digunakan untuk membuktikan pengaruh variabel prediktor terhadap variabel respon secara individu (parsial). Hipotesis yang digunakan adalah.

$$H_0 : \beta_k = 0, k = 0, 1, \dots, j \text{ (Koefisien tidak signifikan)}$$

$$H_1 : \beta_k \neq 0 \text{ (Koefisien signifikan)}$$

Daerah kritis yang digunakan adalah :

$$H_0 \text{ ditolak jika } t_{hitung} > t_{tabel}$$

dengan $t_{tabel} = t_{(n-p-1, \alpha/2)}$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \text{ dimana } 0 \leq R^2 \leq 1 \quad (3.34)$$

3.10. Mean of Square Error (MSE)

Mean of Squares Error (MSE) merupakan salah satu pengukuran kesalahan yang banyak digunakan. Nilai MSE dihitung dengan mengkuadratkan selisih antara nilai

prediksi dengan nilai sebenarnya. Umumnya, semakin kecil nilai MSE maka semakin tinggi tingkat keakuratan nilai suatu prediksi. MSE dapat dihitung dengan menggunakan persamaan berikut (Ghozali, 2009).

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \text{ dimana } i = 1, 2, 3, \dots, n \quad (3.35)$$

3.11. Koefisien Determinasi

Koefisien determinasi (R^2) merupakan suatu nilai atau ukuran yang dapat digunakan untuk mengetahui seberapa jauh kecocokan dari suatu model regresi. Koefisien determinasi mengukur proporsi atau persentase total variasi dalam variabel respon yang dijelaskan oleh model regresi (Gujarati, 2004). Koefisien determinasi didefinisikan sebagai berikut (Sembiring, 1995).

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \text{ dimana } 0 \leq R^2 \leq 1 \quad (3.36)$$

3.12. IPM

Pengukuran pembangunan manusia pertama kali diperkenalkan oleh *United Nations Development Programme* (UNDP) pada tahun 1990. UNDP memperkenalkan sebuah gagasan baru dalam pengukuran pembangunan manusia yang disebut sebagai IPM. Sejak saat itu, IPM dipublikasikan secara berkala dalam laporan tahunan *Human Development Report* (HDR). IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam memperoleh pendapatan, kesehatan, pendidikan dan sebagainya.

Menurut UNDP, Indeks Pembangunan Manusia (IPM) mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup. Sebagai ukuran kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup:

1. Umur panjang dan hidup sehat.
2. Pengetahuan.
3. Standar Hidup Layak.

Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Pada laporan pertamanya, UNDP mengukur dimensi kesehatan dengan menggunakan angka harapan hidup waktu lahir. Selanjutnya untuk mengukur dimensi pengetahuan digunakan angka melek huruf. Adapun untuk mengukur dimensi standar hidup layak digunakan indikator Produk Domestik Bruto (PDRB) per kapita.

Tabel 3.2 Indikator Komponen IPM

| Dimensi | Komponen IPM | Minimum | Maksimum | Keterangan |
|---------------------|-----------------------------------------|-----------------|-------------------|--------------------------------------------|
| Kesehatan | Angka Harapan Hidup saat Lahir (AHH) | 20 | 85 | Sesuai standar global |
| Pengetahuan | Harapan Lama Sekolah (HLS) | 0 | 18 | Sesuai standar global |
| | Rata-rata Lama Sekolah (RLS) | 0 | 15 | Sesuai standar global |
| Standar Hidup Layak | Pengeluaran per kapita Disesuaikan (Rp) | 1.007.436* (Rp) | 26.572.352** (Rp) | UNDP menggunakan PDB riil yang disesuaikan |

Keterangan:

* = Daya beli minimum merupakan garis kemiskinan terendah kabupaten tahun 2010 (data empiris) yaitu di Tolikara-Papua.

** = Daya beli maksimum merupakan nilai tertinggi kabupaten yang diproyeksikan hingga 2025 (akhir RPJPN) yaitu perkiraan pengeluaran per kapita Jakarta Selatan tahun 2025.

Penetapan kategori IPM didasarkan pada skala 0-100 yang terdiri dari:

IPM < 60 = IPM rendah

60 < IPM < 70 = IPM sedang

70 < IPM < 80 = IPM tinggi

IPM > 80 = IPM sangat tinggi

BAB IV

METODOLOGI PENELITIAN

4.1. Data

Data yang digunakan pada penelitian ini adalah berupa data sekunder. Data diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Tengah pada situs <https://jateng.bps.go.id/> yang diakses pada tahun 2017.

4.2. Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Kampus Universitas Islam Indonesia. Penelitian dilakukan pada tanggal 1 Oktober 2017 sampai dengan 10 Februari 2018.

4.3. Variabel dan Definisi Operasional Variabel

Variabel yang digunakan dalam penelitian ini terdiri dari variabel prediktor (variabel bebas) dan variabel respon (variabel terikat).

Tabel 4.1 Definisi Operasional Variabel

| Variabel | Kode | Definisi Operasional Variabel |
|----------|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| IPM | Y | Indeks Pembangunan Manusia mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup |
| AHH | X ₁ | Angka Angka Harapan Hidup saat Lahir didefinisikan sebagai rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. |
| HLS | X ₂ | Angka Harapan Lama Sekolah didefinisikan sebagai lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang. |
| RLS | X ₃ | Rata-rata Lama Sekolah didefinisikan sebagai jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal. |

| Variabel | Kode | Definisi Operasional Variabel |
|----------|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RMG | X ₄ | Rasio murid terhadap guru adalah hasil perbandingan antara jumlah murid dengan jumlah sekolah yang bersangkutan. |
| APS | X ₅ | Angka Partisipasi Sekolah didefinisikan sebagai proporsi dari semua anak yang masih sekolah pada suatu kelompok umur tertentu terhadap penduduk dengan kelompok umur yang sesuai. |
| PDRB | X ₆ | Produk Domestik Regional Bruto didefinisikan sebagai nilai keseluruhan semua barang dan jasa yang diproduksi dalam suatu wilayah dalam suatu jangka waktu tertentu (biasanya satu tahun). |
| TPAK | X ₇ | Tingkat Partisipasi Angkatan Kerja didefinisikan sebagai persentase penduduk usia 15 tahun keatas yang merupakan angkatan kerja. |
| KP | X ₈ | Kepadatan Penduduk didefinisikan sebagai rata-rata jumlah penduduk tiap satu kilometer persegi. Semakin besar angka kepadatan penduduk menunjukkan bahwa semakin padat penduduk yang mendiami wilayah tersebut. |

4.4. Metode Pengumpulan Data

Ada dua metode yang digunakan untuk melakukan pengumpulan data dalam penelitian ini, antara lain:

1. Studi Pustaka

Studi pustaka merupakan langkah awal dalam metode pengumpulan data. Studi pustaka merupakan metode pengumpulan data yang diarahkan kepada pencarian data dan informasi melalui dokumen-dokumen, baik dokumen tertulis, foto-foto, gambar, maupun dokumen elektronik yang dapat mendukung dalam proses penulisan. Hasil penelitian juga akan semakin kredibel apabila didukung foto-foto atau karya tulis akademik dan seni yang telah ada (Sugiyono, 2005). Dalam penelitian ini peneliti mengumpulkan teori relevan sebagai pendukung permasalahan yang akan diteliti melalui buku, jurnal, penelitian sebelumnya dan skripsi.

2. Studi Dokumenter

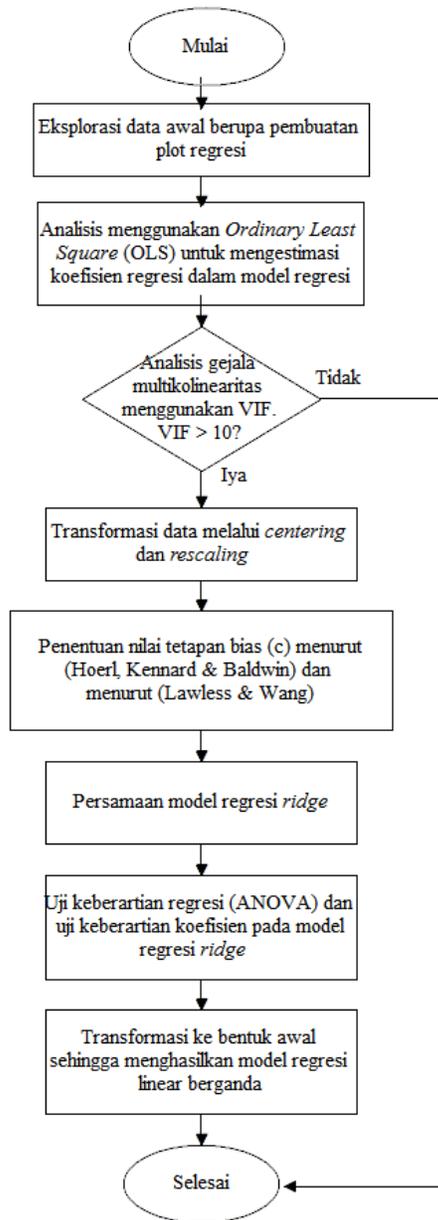
Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari BPS Provinsi Jawa Tengah tentang Indeks Pembangunan Manusia (IPM) tahun 2016.

4.5. Alat dan Cara Organisir Data

Untuk mencapai tujuan penelitian, digunakan bantuan program komputer yaitu program *Excel*, *R* dan *Python*. Langkah-langkah dalam menganalisis data penelitian adalah sebagai berikut:

1. Eksplorasi data awal berupa pembuatan plot untuk melihat apakah terdapat hubungan linier antara variabel prediktor dengan variabel respon.
2. Menganalisis semua data menggunakan metode kuadrat terkecil biasa (*Ordinary Least Square Method*), terutama mendeteksi adanya multikolinearitas antar variabel bebas dengan menggunakan *Variance Inflation Factor (VIF)*. Jika nilai VIF lebih dari 10, maka data mengalami masalah multikolinearitas.
3. Melakukan transformasi data melalui *centering* dan *rescaling*.
4. Penentuan nilai c (tetapan bias) menurut (Hoerl, Kennard & Baldwin, 1975) dan menurut (Lawless & Wang, 1976).
5. Persamaan model regresi *ridge*.
6. Uji *overall* dan uji parsial koefisien pada model regresi *ridge*.
7. Transformasi ke bentuk awal sehingga menghasilkan model regresi linear berganda.

Adapun tahapan penelitian dapat dilihat pada gambar 4.1.

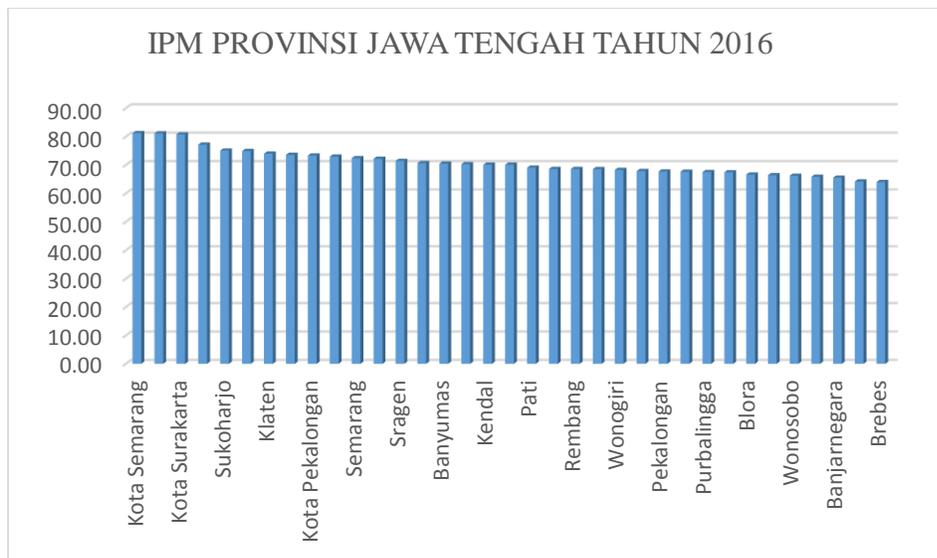


Gambar 4.1. Tahapan Penelitian

BAB V PEMBAHASAN

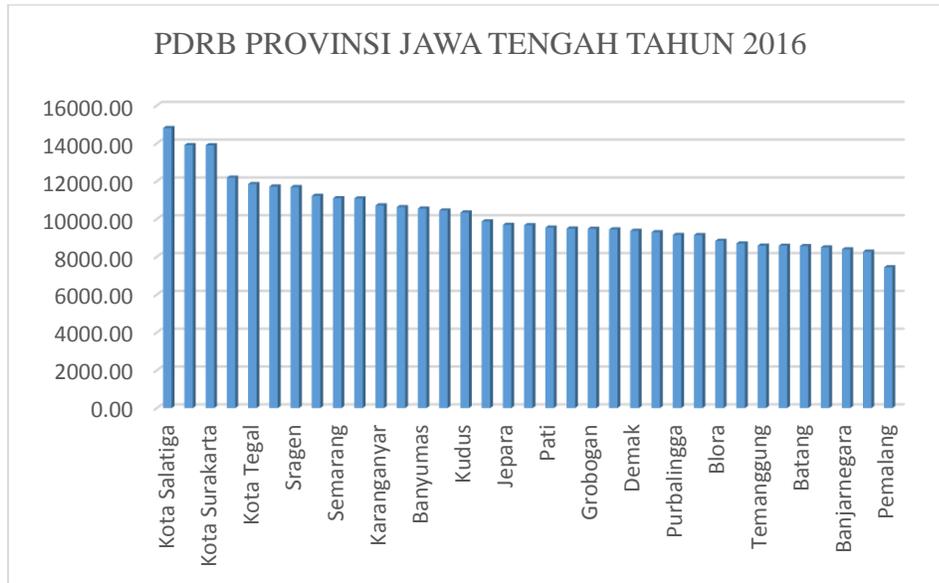
5.1. Analisis Deskriptif

Analisis deskriptif digunakan untuk menggambarkan nilai IPM di Provinsi Jawa Tengah pada tahun 2016. Analisis deskriptif dilakukan dengan bantuan Microsoft Excel, sehingga diperoleh hasil sebagai berikut:



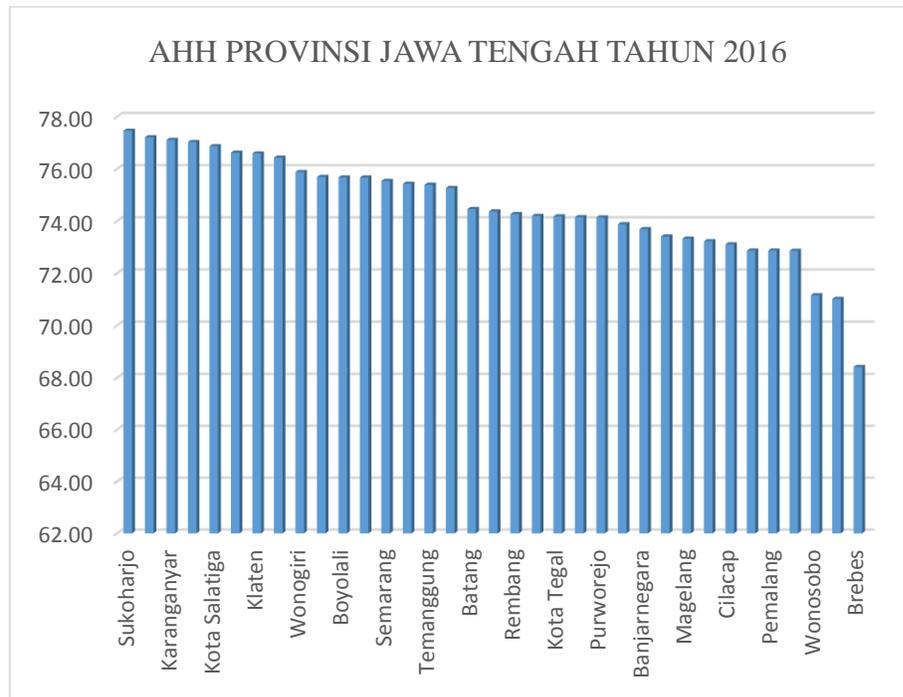
Gambar 5.1 Deskriptif IPM Provinsi Jawa Tengah Tahun 2016

Berdasarkan **Gambar 5.1** di atas capaian pembangunan manusia pada tingkat provinsi menunjukkan variasi antardaerah. Pada tingkat kabupaten/kota, variasi capaian pembangunan manusia justru lebih beragam dibanding pada tingkat provinsi. Capaian pembangunan manusia tertinggi berada di Kota Semarang, sedangkan capaian pembangunan manusia terendah berada di Kabupaten Brebes. Sebagian besar kabupaten/kota telah berada pada status pembangunan manusia “sedang”. Sisanya telah masuk pada kategori “tinggi” dan “rendah. Selain itu, beberapa kabupaten/kota telah masuk pada kategori pembangunan manusia “sangat tinggi”.



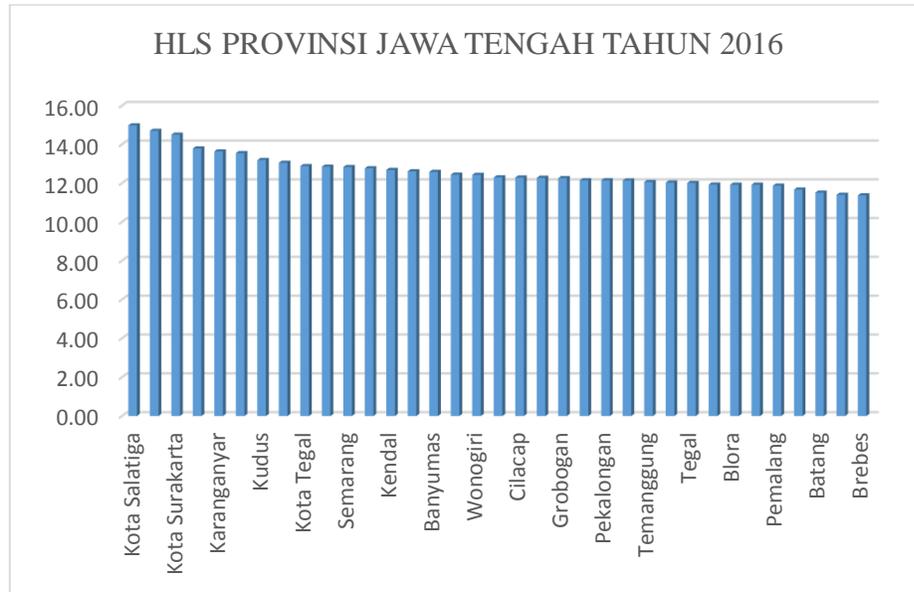
Gambar 5.2 Deskriptif PDRB Provinsi Jawa Tengah Tahun 2016

Berdasarkan **Gambar 5.2** di atas, pengeluaran atau pendapatan telah memberikan sedikit mengenai ukuran pembangunan, seperti yang telah terjadi pada era tahun 70-an. Akan tetapi uang memiliki arti yang penting untuk memperluas pilihan, terutama bagi penduduk miskin. Tahun 2016, pengeluaran per kapita di Provinsi Jawa Tengah telah mencapai 10,1 juta per tahun. Di tingkat kabupaten/kota, pengeluaran per kapita pada tahun 2016 lebih bervariasi. Pengeluaran per kapita berkisar antara 7,5 juta rupiah hingga 15 juta rupiah. Pengeluaran per kapita tertinggi berada di Kota Salatiga, sementara pengeluaran per kapita terendah berada di Kabupaten Pemalang



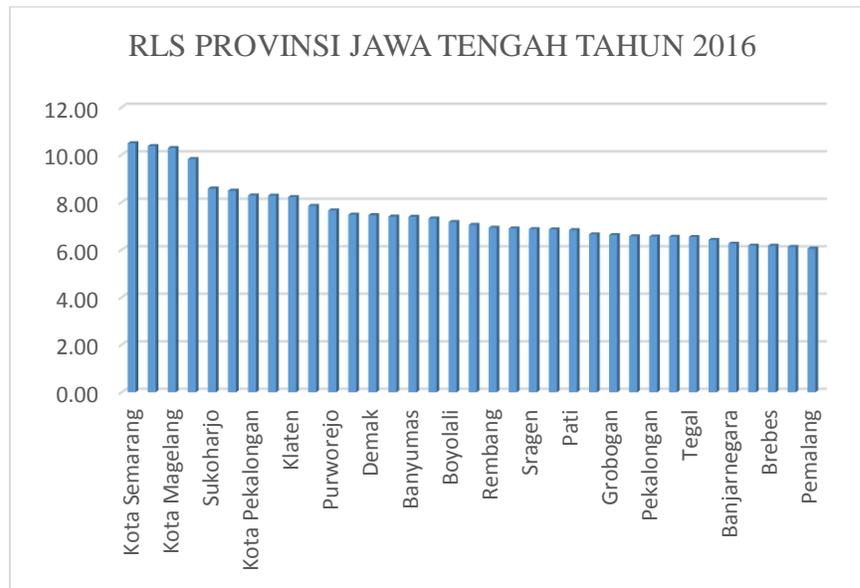
Gambar 5.3 Deskriptif AHH Provinsi Jawa Tengah Tahun 2016

Gambar 5.3 di atas menunjukkan AHH Provinsi Jawa Tengah tahun 2016, selama kurun waku 2010 hingga 2016, AHH di Provinsi Jawa Tengah terus meningkat. Artinya, harapan hidup seorang bayi yang baru lahir untuk dapat hidup lebih lama menjadi semakin tinggi. Pada tahun 2016, AHH saat lahir di Provinsi Jawa Tengah telah mencapai 74,02 tahun. Di tingkat kabupaten/kota, AHH saat lahir pada tahun 2016 berkisar antara 68,41 tahun hingga 77,46 tahun. AHH tertinggi berada di Kabupaten Sukoharjo, sedangkan untuk AHH terendah berada di Kabupaten Brebes.



Gambar 5.4 Deskriptif HLS Provinsi Jawa Tengah Tahun 2016

Gambar 5.4 di atas menunjukkan HLS Jawa Tengah tahun 2016, pendidikan meningkatkan kreativitas dan imajinasi. Sebagai nilai tambah, pendidikan juga akan memperluas pilihan-pilihan lain. Manusia yang berpendidikan akan lebih memperhatikan tingkat kesehatan agar hidup lebih lama. Tidak hanya itu, manusia yang berpendidikan juga akan berpeluang besar mendapatkan pekerjaan dan pendapatan yang lebih layak. Oleh karena itu, pendidikan menjadi penting sebagai sarana untuk meningkatkan kualitas manusia agar dapat memperluas peluang mereka. Selama 6 tahun terakhir, harapan lama sekolah di Provinsi Jawa Tengah terus meningkat. Di tingkat kabupaten/kota harapan lama sekolah di Provinsi Jawa Tengah berkisar antara 11,37 tahun hingga 14,98 tahun. Harapan lama sekolah tertinggi berada di Kota Salatiga, sedangkan untuk harapan lama sekolah terendah berada di Kabupaten Brebes.

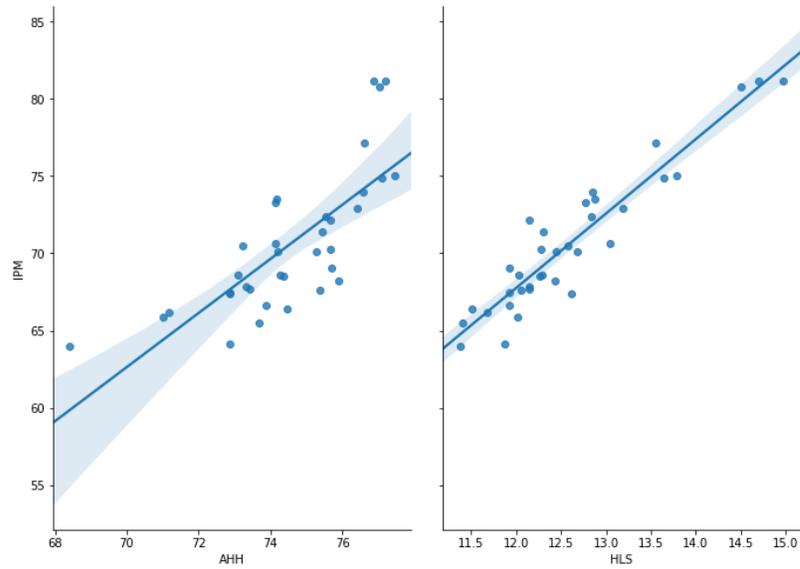


Gambar 5.5 Deskriptif RLS Provinsi Jawa Tengah Tahun 2016

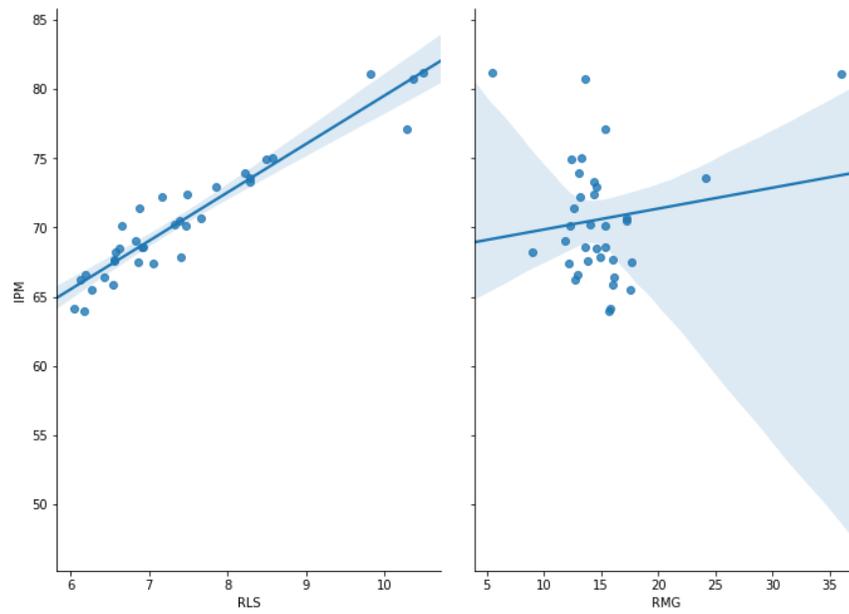
Berdasarkan Gambar 5.5 di atas, RLS penduduk di Provinsi Jawa tengah telah mencapai 7,15 tahun atau setara dengan kelas VII. Capaian dan perkembangan rata-rata lama sekolah di tingkat kabupaten/kota cenderung lebih bervariasi. Di tingkat kabupaten/kota, RLS berkisar antara 6,05 tahun hingga 10,49 tahun. RLS tertinggi berada di Kota Semarang, sedangkan untuk RLS terendah berada di Kabupaten Pemalang.

5.2. Analisis Metode Kuadrat Terkecil (MKT)

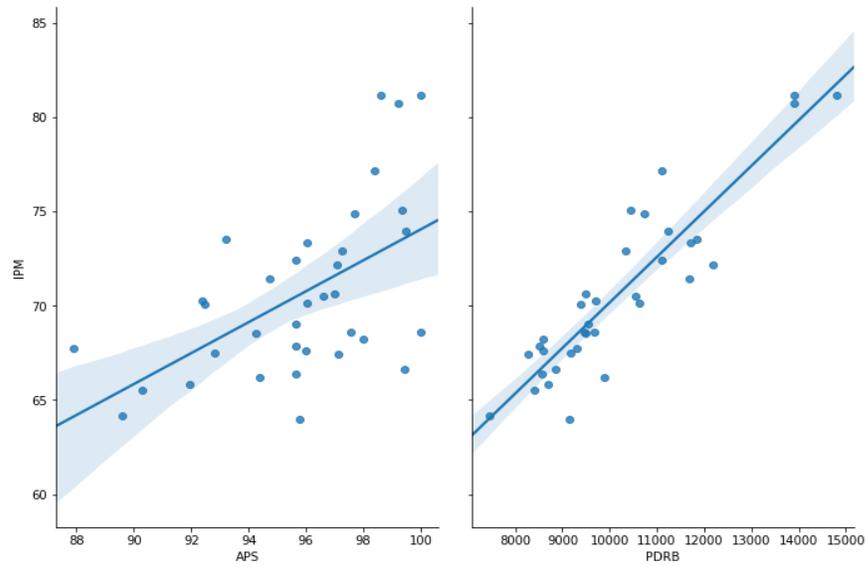
Sebelum dilakukan perhitungan regresi linier, dilakukan eksplorasi awal berupa plot antara variabel prediktor dan variabel respon. Pembuatan plot bertujuan untuk melihat apakah terdapat hubungan linier antara variabel prediktor dengan variabel respon. Hasil plot-nya dapat dilihat pada gambar berikut.



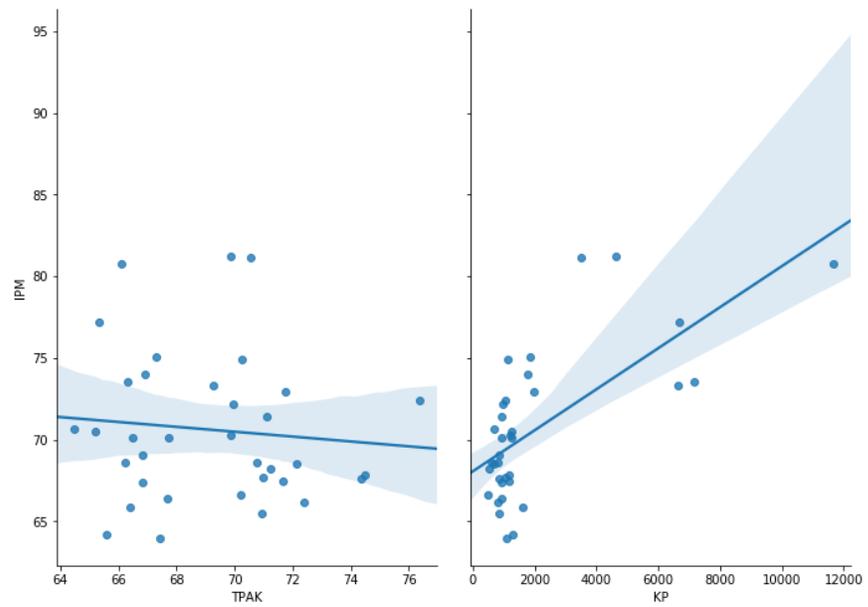
Gambar 5.6 Plot Linier Variabel AHH, HLS dengan IPM



Gambar 5.7 Plot Linier Variabel RLS, RMG dengan IPM



Gambar 5.8 Plot Linier Variabel APS, PDRB dengan IPM



Gambar 5.9 Plot Linier Variabel TPAK, KP dengan IPM

Berdasarkan hasil plot dapat dilihat bahwa pada umumnya pencaran titik cenderung mengikuti garis lurus, sehingga secara grafis dapat disimpulkan bahwa adanya hubungan antara variabel prediktor dan variabel respon. Selanjutnya akan dilakukan analisis MKT untuk melihat pengaruh antara variabel prediktor dengan variabel respon. Hasil analisis MKT yang diperoleh yaitu sebagai berikut.

$$\hat{Y} = 6.059e+00 + 4.442e-01 X_1 + 9.230e-01 X_2 + 1.365e+00 X_3 - 7.729e-03 X_4 + 3.367e-03 X_5 + 9.873e-04 X_6 - 8.111e-03 X_7 - 2.880e-05 X_8.$$

5.3. Uji Overall dan Parsial

Selanjutnya dilakukan pengujian parameter secara *overall* dan pengujian secara parsial. Pengujian parameter secara *overall* atau yang lebih dikenal dengan uji statistik F, pada dasarnya menunjukkan apakah semua variabel prediktor yang dimasukkan dalam model mempunyai pengaruh secara bersama-sama terhadap variabel respon atau untuk mengetahui apakah model regresi dapat digunakan untuk memprediksi hubungan variabel respon atau tidak. Hipotesis untuk pengujian *overall* yang digunakan adalah.
 $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0$, (model regresi tidak signifikan)
 $H_1 : \exists \beta_j \neq 0$, (model regresi signifikan)

Statistik uji yang digunakan dalam pengujian parameter secara *overall* adalah statistik F_{hitung} dengan daerah kritis yaitu:

$$H_0 \text{ ditolak jika } F_{hitung} > F_{tabel}$$

dengan $F_{(8,26,0.05)} = 2.320$.

Hasil pengujian parameter secara *overall* ditampilkan pada tabel berikut.

Tabel 5.1 Hasil Pengujian Parameter Secara *Overall*

| F_{hitung} | F_{tabel} | Kesimpulan |
|--------------|-------------|------------|
| 2181 | 2.320 | Signifikan |

Dari tabel diatas dengan menggunakan tingkat signifikansi sebesar 0,05 maka keputusannya adalah hipotesis H_0 ditolak, artinya model regresi signifikan sehingga

dapat disimpulkan bahwa variabel-variabel prediktor secara *overall* signifikan terhadap variabel respon.

Selanjutnya yaitu dilakukan pengujian secara parsial atau uji statistik t, uji ini pada dasarnya menunjukkan seberapa jauh pengaruh satu variabel prediktor secara individual dalam menerangkan variasi variabel respon. Untuk pengujian parameter secara parsial dengan menggunakan statistik t_{hitung} . Hipotesis yang digunakan adalah:

$H_0 : \beta_k = 0, k = 0,1,2,3,4,5,6,7,8$ (Koefisien tidak signifikan)

$H_1 : \beta_k \neq 0$, (Koefisien signifikan)

Daerah kritis yang digunakan adalah :

H_0 ditolak jika $t_{hitung} > t_{tabel}$

dengan $t_{(26,0.05/2)} = 2.378$

Hasil pengujian parameter secara parsial ditampilkan pada tabel berikut.

Tabel 5.2 Hasil Pengujian Parameter Secara Parsial

| Variabel | Koefisien | t_{hitung} | Kesimpulan |
|----------------|------------|--------------|------------------|
| Konstanta | 6.059e+00 | 2.829 | Signifikan |
| X ₁ | 4.442e-01 | 16.499 | Signifikan |
| X ₂ | 9.230e-01 | 8.034 | Signifikan |
| X ₃ | 1.365e+00 | 14.041 | Signifikan |
| X ₄ | -7.729e-03 | -0.938 | Tidak Signifikan |
| X ₅ | 3.367e-03 | 0.234 | Tidak Signifikan |
| X ₆ | 9.873e-04 | 26.154 | Signifikan |
| X ₇ | -8.111e-03 | -0.628 | Tidak Signifikan |
| X ₈ | -2.880e-05 | -1.123 | Tidak Signifikan |

Berdasarkan tabel diatas dapat disimpulkan bahwa variabel prediktor yang tidak signifikan yaitu X₄,X₅,X₇ dan X₈, hal ini dapat dilihat dari nilai $t_{hitung} < t_{tabel}$. Salah satu cara untuk mendeteksi multikolinearitas adalah melihat nilai koefisien determinasi. Jika koefisien determinasi yang dihasilkan oleh suatu estimasi model regresi cukup tinggi, tetapi secara parsial variabel-variabel prediktor banyak yang tidak signifikan mempengaruhi variabel respon maka hal tersebut mengindikasikan adanya multikolinearitas (Ghozali, 2013). Hasil analisis menunjukkan bahwa nilai *adjusted R*² yaitu 99,8% dan secara uji *overall* variabel prediktor secara bersama-sama berpengaruh

terhadap variabel respon, akan tetapi secara uji parsial banyak variabel prediktor tidak signifikan mempengaruhi model.

5.4. Uji Multikolinearitas

Dalam mendeteksi multikolinearitas pada regresi linier berganda menggunakan ukuran nilai VIF. Apabila nilai VIF lebih besar dari 10 mengindikasikan bahwa ada masalah multikolinearitas yang serius. Nilai VIF diperoleh dengan cara meregresikan variabel X_p dengan variabel-variabel prediktor lainnya yang bertujuan untuk mengukur kombinasi pengaruh ketergantungan antara variabel-variabel prediktor tersebut. Nilai VIF untuk masing-masing variabel ditampilkan sebagai berikut.

Tabel 5.3 Tabel Multikolinearitas

| Variabel | Nilai VIF |
|----------------|-----------|
| AHH (X_1) | 2.403615 |
| HLS (X_2) | 8.835969 |
| RLS (X_3) | 12.246497 |
| RMG (X_4) | 1.262296 |
| APS (X_5) | 1.600125 |
| PDRB (X_6) | 3.515975 |
| TPAK(X_7) | 1.2282212 |
| KP(X_8) | 3.330172 |

Berdasarkan tabel di atas dapat dilihat untuk variabel X_3 memiliki nilai VIF lebih besar dari 10 sehingga dengan menggunakan tingkat signifikansi sebesar 0,05 dapat disimpulkan bahwa pada data tersebut terdapat permasalahan multikolinearitas pada variabel-variabel prediktor sehingga perlu diatasi dengan menggunakan metode tertentu, salah satunya adalah metode regresi *ridge*.

5.5. Regresi Ridge

Sebelum pemodelan regresi *ridge* dibentuk, perlu dilakukan transformasi data yang disebut dengan pemusatan dan penskalaan (*centering & scaling*) untuk meminimumkan kesalahan dalam pembulatan data dan juga prosedur ini akan mengakibatkan hilangnya β_0 yang membuat perhitungan untuk mencari model regresi menjadi lebih sederhana dan lebih mudah. Hasil dari transformasi ditampilkan pada **Tabel 5.4** sebagai berikut:

Tabel 5.4 Data Transformasi

| IPM* | AHH* | HLS* | RLS* | RMG* | APS* | PDRB* | TPAK* | KP* |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -0.449 | -0.742 | -0.345 | -0.450 | 0.094 | 0.586 | -0.300 | -1.016 | -0.512 |
| -0.027 | -0.681 | -0.014 | -0.048 | 0.482 | 0.270 | 0.222 | -1.371 | -0.327 |
| -0.700 | -0.871 | -0.757 | -0.483 | 0.593 | -1.003 | -0.608 | 0.863 | -0.358 |
| -1.138 | -0.445 | -1.362 | -0.976 | 0.563 | -1.853 | -1.060 | 0.612 | -0.491 |
| -0.715 | -0.866 | 0.020 | -0.327 | -0.601 | 0.445 | -1.134 | -0.803 | -0.458 |
| 0.011 | -0.214 | 0.523 | 0.174 | 0.482 | 0.401 | -0.407 | -1.615 | -0.557 |
| -0.988 | -1.743 | -1.054 | -1.091 | -0.495 | -0.482 | -0.181 | 1.101 | -0.514 |
| -0.617 | -0.630 | -0.505 | -0.040 | -0.010 | -0.052 | -1.000 | 1.831 | -0.362 |
| 0.351 | 0.572 | -0.517 | -0.229 | -0.383 | 0.425 | 1.197 | 0.271 | -0.447 |
| 0.751 | 1.044 | 0.294 | 0.634 | -0.426 | 1.238 | 0.622 | -0.772 | -0.106 |
| 0.995 | 1.491 | 1.368 | 0.929 | -0.376 | 1.191 | 0.161 | -0.648 | -0.067 |
| -0.532 | 0.680 | -0.185 | -0.721 | -1.306 | 0.734 | -0.947 | 0.705 | -0.626 |
| 0.959 | 1.311 | 1.197 | 0.855 | -0.554 | 0.633 | 0.322 | 0.368 | -0.378 |
| 0.183 | 0.449 | -0.334 | -0.475 | -0.519 | -0.368 | 0.897 | 0.671 | -0.456 |
| -0.467 | -0.096 | -0.380 | -0.680 | -0.090 | -0.519 | -0.413 | 1.025 | -0.558 |
| -0.894 | -0.347 | -0.768 | -1.042 | -0.432 | 1.221 | -0.795 | 0.357 | -0.644 |
| -0.449 | -0.147 | -0.642 | -0.426 | -0.296 | 1.409 | -0.433 | 0.554 | -0.588 |
| -0.353 | 0.582 | -0.768 | -0.508 | -0.680 | -0.055 | -0.377 | -0.806 | -0.498 |
| 0.521 | 0.962 | 0.683 | 0.330 | -0.090 | 0.489 | 0.099 | 0.887 | -0.021 |
| -0.080 | 0.572 | -0.357 | -0.105 | -0.209 | -1.144 | -0.289 | 0.233 | -0.344 |
| -0.114 | 0.366 | -0.174 | 0.010 | -0.584 | -1.127 | -0.479 | -0.496 | -0.320 |
| 0.400 | 0.505 | 0.272 | 0.026 | -0.129 | -0.052 | 0.548 | 2.478 | -0.398 |
| -0.673 | 0.428 | -0.608 | -0.738 | -0.246 | 0.069 | -0.945 | 1.789 | -0.484 |
| -0.112 | -0.183 | 0.100 | -0.656 | 0.077 | 0.079 | 0.268 | -0.923 | -0.450 |
| -0.946 | -0.049 | -1.236 | -0.844 | 0.244 | -0.059 | -0.960 | -0.507 | -0.448 |
| -0.648 | -0.588 | -0.505 | -0.729 | 0.224 | -2.656 | -0.524 | 0.622 | -0.406 |

| IPM* | AHH* | HLS* | RLS* | RMG* | APS* | PDRB* | TPAK* | KP* |
|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|------------|
| -1.440 | -0.866 | -0.825 | -1.148 | 0.168 | -2.082 | -1.627 | -1.240 | -0.313 |
| -1.067 | -1.815 | -0.665 | -0.746 | 0.220 | -1.295 | -0.876 | -0.951 | -0.168 |
| -1.482 | -3.155 | -1.396 | -1.050 | 0.149 | -0.005 | -0.615 | -0.603 | -0.395 |
| 1.465 | 1.059 | 1.094 | 2.334 | 0.086 | 0.862 | 0.541 | -1.326 | 1.934 |
| 2.270 | 1.270 | 2.180 | 2.399 | -0.300 | 1.147 | 2.213 | -1.058 | 4.008 |
| 2.355 | 1.188 | 2.728 | 1.948 | 4.560 | 1.409 | 2.755 | 0.467 | 0.619 |
| 2.366 | 1.362 | 2.408 | 2.498 | -2.063 | 0.936 | 2.218 | 0.240 | 1.079 |
| 0.606 | -0.209 | 0.203 | 0.691 | -0.131 | 0.079 | 0.916 | 0.037 | 1.922 |
| 0.658 | -0.193 | 0.329 | 0.683 | 1.976 | -0.872 | 0.992 | -0.978 | 2.135 |

Keterangan:

* = Data yang telah dipusatkan dan diskalakan.

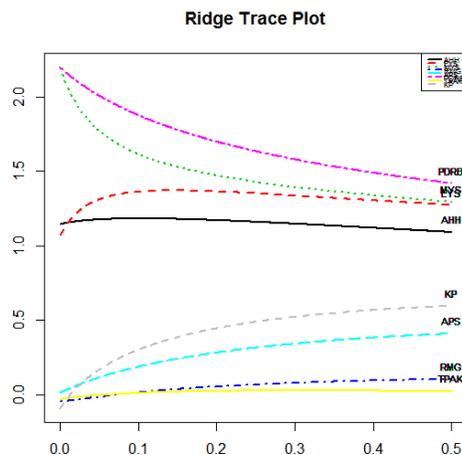
Apabila data telah dipusatkan dan diskalakan, maka dilanjutkan penentuan nilai tetapan bias yang merupakan hal paling penting dalam melakukan analisis regresi *ridge*. Penentuan nilai tetapan bias dapat menggunakan tabel VIF. Nilai dari koefisien VIF $\hat{\beta}(c)$ dengan berbagai kemungkinan tetapan bias dapat dilihat pada tabel berikut.

Tabel 5.5 Nilai VIF $\hat{\beta}(c)$ dengan Berbagai Nilai c

| | AHH* | HLS* | RLS* | RMG* | APS* | PDRB* | TPAK* | KP* |
|---------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|------------|
| c=0 | 2.404 | 8.836 | 12.246 | 1.262 | 1.600 | 3.516 | 1.228 | 3.330 |
| c=0.012 | 2.208 | 6.357 | 8.314 | 1.200 | 1.519 | 3.162 | 1.178 | 2.758 |
| c=0.034 | 1.910 | 4.019 | 4.780 | 1.107 | 1.387 | 2.633 | 1.095 | 2.145 |
| c=0.04 | 1.852 | 3.684 | 4.298 | 1.089 | 1.360 | 2.530 | 1.078 | 2.046 |
| c=0.08 | 1.495 | 2.189 | 2.269 | 0.967 | 1.181 | 1.918 | 0.963 | 1.537 |
| c=0.12 | 1.244 | 1.504 | 1.438 | 0.872 | 1.041 | 1.511 | 0.871 | 1.238 |
| c=0.16 | 1.058 | 1.117 | 1.009 | 0.794 | 0.928 | 1.225 | 0.794 | 1.036 |
| c=0.2 | 0.916 | 0.872 | 0.757 | 0.728 | 0.835 | 1.016 | 0.729 | 0.888 |
| c=0.24 | 0.804 | 0.705 | 0.594 | 0.672 | 0.756 | 0.859 | 0.674 | 0.774 |
| c=0.28 | 0.713 | 0.586 | 0.483 | 0.623 | 0.689 | 0.737 | 0.625 | 0.684 |
| c=0.32 | 0.638 | 0.496 | 0.403 | 0.581 | 0.631 | 0.641 | 0.582 | 0.611 |
| c=0.36 | 0.576 | 0.428 | 0.343 | 0.542 | 0.581 | 0.563 | 0.544 | 0.551 |
| c=0.4 | 0.523 | 0.374 | 0.297 | 0.508 | 0.537 | 0.500 | 0.510 | 0.500 |
| c=0.44 | 0.478 | 0.330 | 0.261 | 0.478 | 0.498 | 0.447 | 0.479 | 0.457 |
| c=0.48 | 0.439 | 0.295 | 0.232 | 0.450 | 0.464 | 0.403 | 0.451 | 0.420 |
| c=0.52 | 0.405 | 0.265 | 0.208 | 0.425 | 0.433 | 0.366 | 0.426 | 0.387 |

| | AHH* | HLS* | RLS* | RMG* | APS* | PDRB* | TPAK* | KP* |
|--------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|------------|
| c=0.56 | 0.375 | 0.241 | 0.189 | 0.402 | 0.405 | 0.334 | 0.403 | 0.359 |
| c=0.6 | 0.349 | 0.220 | 0.172 | 0.381 | 0.380 | 0.306 | 0.381 | 0.334 |
| c=0.64 | 0.326 | 0.202 | 0.158 | 0.361 | 0.358 | 0.282 | 0.362 | 0.312 |
| c=0.68 | 0.305 | 0.186 | 0.146 | 0.343 | 0.337 | 0.261 | 0.344 | 0.292 |
| c=0.72 | 0.286 | 0.172 | 0.136 | 0.327 | 0.319 | 0.242 | 0.328 | 0.274 |
| c=0.76 | 0.269 | 0.161 | 0.127 | 0.312 | 0.301 | 0.226 | 0.312 | 0.258 |
| c=0.8 | 0.254 | 0.150 | 0.119 | 0.297 | 0.286 | 0.211 | 0.298 | 0.244 |
| c=0.84 | 0.240 | 0.141 | 0.112 | 0.284 | 0.271 | 0.198 | 0.285 | 0.230 |
| c=0.88 | 0.227 | 0.132 | 0.106 | 0.272 | 0.258 | 0.186 | 0.272 | 0.218 |
| c=0.92 | 0.215 | 0.125 | 0.100 | 0.260 | 0.246 | 0.175 | 0.261 | 0.207 |
| c=0.96 | 0.205 | 0.118 | 0.095 | 0.250 | 0.234 | 0.166 | 0.250 | 0.197 |
| c=1 | 0.195 | 0.112 | 0.090 | 0.239 | 0.224 | 0.157 | 0.240 | 0.188 |
| c=1.04 | 0.186 | 0.107 | 0.086 | 0.230 | 0.214 | 0.149 | 0.230 | 0.179 |
| c=1.08 | 0.177 | 0.101 | 0.082 | 0.221 | 0.205 | 0.142 | 0.221 | 0.171 |
| c=1.12 | 0.169 | 0.097 | 0.079 | 0.213 | 0.196 | 0.135 | 0.213 | 0.163 |
| c=1.16 | 0.162 | 0.093 | 0.076 | 0.205 | 0.188 | 0.129 | 0.205 | 0.157 |
| c=1.2 | 0.155 | 0.089 | 0.073 | 0.197 | 0.180 | 0.123 | 0.198 | 0.150 |
| c=1.24 | 0.149 | 0.085 | 0.070 | 0.190 | 0.173 | 0.118 | 0.191 | 0.144 |
| c=1.28 | 0.143 | 0.082 | 0.067 | 0.184 | 0.167 | 0.113 | 0.184 | 0.138 |
| c=1.32 | 0.138 | 0.079 | 0.065 | 0.177 | 0.160 | 0.108 | 0.177 | 0.133 |
| c=1.36 | 0.132 | 0.076 | 0.063 | 0.171 | 0.155 | 0.104 | 0.171 | 0.128 |
| c=1.4 | 0.128 | 0.073 | 0.061 | 0.166 | 0.149 | 0.100 | 0.166 | 0.123 |
| c=1.44 | 0.123 | 0.070 | 0.059 | 0.160 | 0.144 | 0.096 | 0.160 | 0.119 |
| c=1.48 | 0.119 | 0.068 | 0.057 | 0.155 | 0.139 | 0.092 | 0.155 | 0.115 |
| c=1.52 | 0.115 | 0.066 | 0.055 | 0.150 | 0.134 | 0.089 | 0.150 | 0.111 |
| c=1.56 | 0.111 | 0.064 | 0.054 | 0.145 | 0.130 | 0.086 | 0.146 | 0.107 |
| c=1.6 | 0.107 | 0.062 | 0.052 | 0.141 | 0.125 | 0.083 | 0.141 | 0.104 |
| c=1.64 | 0.104 | 0.060 | 0.051 | 0.137 | 0.121 | 0.080 | 0.137 | 0.100 |
| c=1.68 | 0.100 | 0.058 | 0.049 | 0.133 | 0.118 | 0.078 | 0.133 | 0.097 |
| c=1.72 | 0.097 | 0.056 | 0.048 | 0.129 | 0.114 | 0.075 | 0.129 | 0.094 |
| c=1.76 | 0.094 | 0.055 | 0.047 | 0.125 | 0.110 | 0.073 | 0.125 | 0.091 |
| c=1.8 | 0.091 | 0.053 | 0.046 | 0.122 | 0.107 | 0.071 | 0.122 | 0.089 |
| c=1.84 | 0.089 | 0.052 | 0.045 | 0.118 | 0.104 | 0.069 | 0.118 | 0.086 |
| c=1.88 | 0.086 | 0.051 | 0.043 | 0.115 | 0.101 | 0.067 | 0.115 | 0.084 |
| c=1.92 | 0.084 | 0.049 | 0.042 | 0.112 | 0.098 | 0.065 | 0.112 | 0.081 |
| c=1.96 | 0.081 | 0.048 | 0.042 | 0.109 | 0.095 | 0.063 | 0.109 | 0.079 |
| c=2 | 0.079 | 0.047 | 0.041 | 0.106 | 0.093 | 0.061 | 0.106 | 0.077 |

Dari tabel 5.6 di atas tampak bahwa mulai tetapan bias $c = 0,00$ sampai dengan nilai $c = 2,00$ VIF koefisien estimator $\hat{\beta}(c)$ semakin menyusut mendekati nol. Selain dengan menggunakan nilai VIF, pemilihan nilai tetapan bias dapat dilakukan berdasarkan pada pola kecenderungan jejak *ridge* atau *ridge trace* yang menghasilkan koefisien estimator yang relatif stabil.



Gambar 5.10 *Ridge Trace*

Hasil *ridge trace* pada **Gambar 5.10** menunjukkan hasil yang bersifat subjektif dalam pemilihan nilai tetapan bias. Hal tersebut dikarenakan sulitnya menentukan nilai tetapan bias c yang paling minimum ketika nilai $\beta(c)$ mulai stabil pada setiap peubah bebas. Penentuan nilai tetapan bias dapat menggunakan beberapa metode, dalam penelitian ini menggunakan dua metode yaitu metode (Hoerl, Kennard & Baldwin, 1975) dan metode (Lawless & Wang, 1976).

Tabel 5.6 Nilai Tetapan Bias

| Metode | Nilai c |
|----------------------------------|------------|
| (Hoerl, Kennard & Baldwin, 1975) | 0.03471074 |
| (Lawless & Wang, 1976) | 0.01203737 |

Berdasarkan nilai tetapan bias yang dipilih, maka penduga koefisien hasil analisis dapat dilihat pada **Tabel 5.7**. Adanya nilai tetapan bias pada regresi *ridge* menyebabkan

dugaan koefisien regresi yang dihasilkan semakin menyusut. Dugaan koefisien regresi *ridge* cenderung lebih kecil dibandingkan dengan dugaan koefisien regresi MKT. Pemilihan nilai tetapan bias berdasarkan pertimbangan pada koefisien estimator yang relatif stabil.

Tabel 5.7 Koefisien Regresi Hasil MKT dan Regresi *Ridge*

| Peubah | MKT | Metode (Hoerl, Kennard & Baldwin, 1975) | Metode (Lawless & Wang, 1976) |
|----------------|------------|-----------------------------------------|-------------------------------|
| Intersep | 6.059e+00 | -1.621e-10 | -2.057e-10 |
| X ₁ | 4.442e-01 | 1.978e-01 | 1.954e-01 |
| X ₂ | 9.230e-01 | 2.153e-01 | 1.978e-01 |
| X ₃ | 1.365e+00 | 3.128e-01 | 3.443e-01 |
| X ₄ | -7.729e-03 | -4.260e-03 | -6.817e-03 |
| X ₅ | 3.367e-03 | 1.498e-02 | 7.137e-03 |
| X ₆ | 9.873e-04 | 3.475e-01 | 3.620e-01 |
| X ₇ | -8.111e-03 | -1.944e-03 | -3.964e-03 |
| X ₈ | -2.880e-05 | 1.881e-02 | -7.824e-04 |

Keragaman koefisien regresi hasil analisis dengan menggunakan MKT dan regresi *ridge* dapat terlihat dari nilai variansi koefisien regresi. Pada **Tabel 5.8** terlihat nilai variansi, bias, MSE, koefisien determinasi (R^2) dan *adjusted R²* antara MKT dan regresi *ridge*.

Tabel 5.8 Perbandingan Nilai Variansi, Bias, MSE, R^2 dan *Adjusted R²*

| Metode | Variansi | Bias | MSE | R^2 | <i>Adjusted R²</i> |
|-----------------------------------------|----------|--------|--------|--------|-------------------------------|
| MKT | 0.0664 | 0.000 | 0.0664 | 0.9985 | 0.9981 |
| Metode (Hoerl, Kennard & Baldwin, 1975) | 0.0514 | 0.2312 | 0.2826 | 0.9768 | 0.9707 |
| Metode (Lawless & Wang, 1976) | 0.0551 | 0.0482 | 0.1033 | 0.9905 | 0.9881 |

Pada **Tabel 5.8** dapat dilihat bahwa keragaman yang dihasilkan oleh kedua metode penentuan nilai tetapan bias menurut (Hoerl, Kennard & Baldwin, 1975) dan menurut (Lawless & Wang, 1976) relatif lebih kecil dibandingkan dengan MKT.

Selanjutnya dari koefisien estimator yang didapat akan terbentuk persamaan regresi *ridge* untuk masing-masing metode penentuan nilai tetapan bias sebagai berikut:

Metode (Hoerl, Kennard & Baldwin, 1975):

$$\hat{Y}^R = 1.16998342 X_1 + 1.27361831 X_2 + 1.85074498 X_3 - 0.02520362 X_4 + 0.08860856 X_5 + 2.05582870 X_6 - 0.01150329 X_7 - 0.11126214 X_8.$$

Metode (Lawless & Wang, 1976) :

$$\hat{Y}^R = 1.15627353 X_1 + 1.17024967 X_2 + 2.03688026 X_3 - 0.04033228 X_4 + 0.04222384 X_5 + 2.14146981 X_6 - 0.02344904 X_7 - 0.00462853 X_8.$$

Dari persamaan model regresi *ridge* tersebut dilakukan pengujian keberartian regresi dan koefisien dengan tingkat signifikansi sebesar 5% sebagai berikut:

Tabel 5.9 Uji *Overall* Regresi *Ridge*

| Metode | F_{hitung} | Kesimpulan |
|-----------------------------------------|--------------|------------|
| Metode (Hoerl, Kennard & Baldwin, 1975) | 1620.707 | Signifikan |
| Metode (Lawless & Wang, 1976) | 2115.069 | Signifikan |

Berdasarkan pada tabel diatas dapat diketahui bahwa kesimpulan tolak H_0 untuk semua metode, artinya variabel prediktor untuk metode penentuan nilai tetapan bias menurut (Hoerl, Kennard & Baldwin, 1975) dan menurut (Lawless & Wang, 1976) secara *overall* signifikan di dalam model.

Selanjutnya yaitu dilakukan pengujian secara parsial untuk masing-masing metode yang digunakan. Untuk metode (Hoerl, Kennard & Baldwin, 1975), pengujian parameter secara parsial dengan menggunakan statistik t_{hitung} . Hipotesis yang digunakan adalah:

$$H_0 : \beta_k = 0, k = 1,2,3,4,5,6,7,8 \text{ (Koefisien tidak signifikan)}$$

$$H_1 : \beta_k \neq 0, \text{ (Koefisien signifikan)}$$

Daerah kritis yang digunakan adalah :

$$H_0 \text{ ditolak jika } t_{hitung} > t_{tabel}$$

$$\text{dengan } t_{(26,0.05/2)} = 2.378$$

Hasil pengujian parameter secara parsial metode (Hoerl, Kennard & Baldwin, 1975) ditampilkan pada tabel berikut.

Tabel 5.10 Hasil Pengujian Parameter Secara Parsial Metode (Hoerl, Kennard & Baldwin, 1975)

| Variabel | Koefisien | t_{hitung} | Kesimpulan |
|----------------|------------|--------------|------------------|
| X ₁ | 1.978e-01 | 16.302 | Signifikan |
| X ₂ | 2.153e-01 | 12.237 | Signifikan |
| X ₃ | 3.128e-01 | 16.303 | Signifikan |
| X ₄ | -4.260e-03 | -0.461 | Tidak Signifikan |
| X ₅ | 1.498e-02 | 1.449 | Tidak Signifikan |
| X ₆ | 3.475e-01 | 24.403 | Signifikan |
| X ₇ | -1.944e-03 | -0.212 | Tidak Signifikan |
| X ₈ | 1.881e-02 | 1.463 | Tidak Signifikan |

Berdasarkan pada **Tabel 5.0** di atas dapat dilihat dengan menggunakan tingkat signifikansi sebesar 0,05 bahwa variabel X₄, X₅, X₇ dan X₈ tidak signifikan. Sedangkan variabel prediktor yang terdiri dari variabel konstanta, X₁, X₂, X₃ dan X₆ secara parsial signifikan terhadap variabel respon.

Selanjutnya untuk metode (Lawless & Wang, 1976), pengujian parameter secara parsial dengan menggunakan statistik t_{hitung} . Hipotesis yang digunakan adalah:

$$H_0 : \beta_k = 0, k = 1,2,3,4,5,6,7,8 \text{ (Koefisien tidak signifikan)}$$

$$H_1 : \beta_k \neq 0, \text{ (Koefisien signifikan)}$$

Daerah kritis yang digunakan adalah :

$$H_0 \text{ ditolak jika } t_{hitung} > t_{tabel}$$

$$\text{dengan } t_{(26,0.05/2)} = 2.378$$

Hasil pengujian parameter secara parsial metode (Lawless & Wang, 1976) ditampilkan pada tabel berikut.

Tabel 5.11 Hasil Pengujian Parameter Secara Parsial Metode (Lawless & Wang, 1976)

| Variabel | Koefisien | t_{hitung} | Kesimpulan |
|----------------|------------|--------------|------------------|
| X ₁ | 1.954e-01 | 17.122 | Signifikan |
| X ₂ | 1.978e-01 | 10.213 | Signifikan |
| X ₃ | 3.443e-01 | 15.544 | Signifikan |
| X ₄ | -6.817e-03 | -0.810 | Tidak Signifikan |
| X ₅ | 7.137e-03 | 0.754 | Tidak Signifikan |
| X ₆ | 3.620e-01 | 26.497 | Signifikan |
| X ₇ | -3.964e-03 | -0.475 | Tidak Signifikan |
| X ₈ | -7.824e-04 | -0.061 | Tidak Signifikan |

Berdasarkan pada **Tabel 5.11** di atas dapat dilihat dengan menggunakan tingkat signifikansi sebesar 0,05 bahwa variabel prediktor yang terdiri dari variabel X₄, X₅, X₇ dan X₈ secara parsial tidak signifikan terhadap variabel respon, sedangkan sisanya yaitu variabel konstanta, X₁, X₂, X₃ dan X₆ secara parsial signifikan terhadap variabel respon.

Setelah dilakukan uji parsial regresi *ridge*, maka didapatkan permodelan regresi berganda untuk masing-masing metode sebagai berikut:

Metode (Hoerl, Kennard & Baldwin, 1975):

$$\hat{Y} = -328.636 + 2.685789 X_1 + 6.506847 X_2 + 6.796776 X_3 + 0.005471 X_6.$$

Metode (Lawless & Wang, 1976):

$$\hat{Y} = -319.284 + 2.65431686 X_1 + 5.978742361 X_2 + 7.4803488 X_3 + 0.005698536 X_6.$$

Berdasarkan hasil analisis diatas untuk menentukan nilai tetapan bias maka peneliti menggunakan metode (Lawless & Wang, 1976) karena memiliki nilai bias dan MSE yang lebih kecil serta nilai *adjusted R²* yang lebih besar dibandingkan metode (Hoerl, Kennard & Baldwin, 1975). Berdasarkan perbandingan dua metode penentuan nilai bias dapat dikatakan bahwa metode penentuan nilai tetapan bias menurut (Lawless & Wang, 1976) merupakan metode terbaik dalam mengatasi masalah multikolinearitas pada kasus IPM di Provinsi Jawa Tengah tahun 2016.

BAB VI

PENUTUP

6.1. Kesimpulan

Berdasarkan hasil analisis dan pembahasan, maka dapat diperoleh kesimpulan sebagai berikut:

1. Penentuan nilai bias regresi *ridge* menggunakan metode (Hoerl, Kennard & Baldwin, 1975) dan metode (Lawless & Wang, 1976) diperoleh hasil bahwa metode (Lawless & Wang, 1976) memiliki nilai bias dan MSE yang lebih kecil serta nilai *adjusted R²* yang lebih besar dibandingkan metode (Hoerl, Kennard & Baldwin, 1975). Berdasarkan perbandingan dua metode penentuan nilai bias dapat dikatakan bahwa metode penentuan nilai tetapan bias menurut (Lawless & Wang, 1976) merupakan metode terbaik dalam mengatasi masalah multikolinearitas pada kasus indeks pembangunan manusia di Provinsi Jawa Tengah tahun 2016.
2. Penerapan metode regresi *ridge* pada skripsi ini diambil dari kasus IPM di Provinsi Jawa Tengah tahun 2016. Penentuan nilai tetapan bias regresi *ridge* menggunakan metode (Lawless & Wang, 1976) diperoleh nilai tetapan bias c yaitu sebesar 0.01203737. Nilai ini menunjukkan koefisien $\hat{\beta}$ lebih stabil dibandingkan dengan metode (Hoerl, Kennard & Baldwin, 1975), sehingga diperoleh persamaan regresi *ridge* untuk metode (Lawless & Wang, 1976) yaitu $\hat{Y} = -319.284 + 2.65431686 X_1 + 5.978742361 X_2 + 7.4803488 X_3 + 0.005698536 X_6$.

6.2. Saran

Adapun saran yang dapat penulis berikan adalah sebagai berikut:

1. Penanganan masalah multikolinearitas pada skripsi ini menggunakan metode regresi *ridge*, untuk penelitian selanjutnya dapat menggunakan metode lain untuk mengatasi multikolinearitas dan selanjutnya dapat dibandingkan dengan metode regresi *ridge* untuk memperoleh hasil terbaik.
2. Untuk model regresi yang mengalami masalah multikolinearitas, sebaiknya peneliti tidak mengeluarkan variabel prediktor yang berkorelasi tinggi pada model untuk penanganan multikolinearitas karena akan menyebabkan interpretasi hasil analisis jauh dari fakta.

DAFTAR PUSTAKA

- Astuti, A.D. 2014. *Partial Least Square (PLS) dan Principal Component Regression (PCR) untuk Regresi Linier dengan Multikolinearitas pada Kasus Indeks Pembangunan Manusia di Kabupaten Gunung Kidul*. Universitas Negeri Yogyakarta.
- Ayunanda, M., & Ismaini, Z. 2013. Analisis Statistika Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur dengan Menggunakan Regresi Panel. *Jurnal Sains dan Seni Pomits*, Vol. 2, No. 2, 2337-3520.
- Badan Pusat Statistik Provinsi Jawa Tengah (BPS), diakses dari <https://jateng.bps.go.id/>, diakses pada tanggal 20 Oktober 2017 pada jam 19.30 WIB.
- Bain, L.J. 1992. *Introduction to Probability and Mathematical Statistics*. Jakarta: Gramedia Pustaka.
- Dereny, M. El. & Rashwan, N.I. 2011. *Solving Multicollinearity Problem Using Ridge Regression Models*. *Int. J. Contemp. Math. Sciences* Vol. 6 No. 12 Hal. 585-600.
- Draper, N.R. & Smith, H. 1992. *Analisis Regresi Terapan*. Edisi 2. (Terjemahan: Bambang-Sumantri). Jakarta. PT. Gramedia Pustaka Utama
- Ghozali, I. 2009. *Aplikasi Analisis Multivariate Dengan Program SPSS*. Edisi Keempat. Semarang: Universitas Diponegoro.
- Ghozali, I. 2013. *Statistik Nonparametrik*. Semarang: Badan Penerbit UNDIP.
- Hastie, T., Robert, T., & Jerome, F. 2008. *The Elements of Statistical Learning. Data Mining, Inference, and prediction*. Edisi Kedua. New York: Spring.
- Hoerl, A.E., & R.W. Kennard. 1970. *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. <http://statgen.ucr.edu/file/STAT288/hoerl70a.pdf> (Diakses 21 Oktober 2017).

- Hoerl, A.E, Kennard, R.W., & Baldwin, K.F. 1975. A Simulation Study of Ridge and Other Regression Estimators. *Communication in Statistics*, 5, 308-323.
- Kartika A., L., Maria, B., & Rahma, F. 2013. Pendekatan Partial Least Square Regression untuk Mengatasi Multikolinearitas dalam Regresi Logistik Ordinal. *Jurna MIPA Universitas Brawijaya*.
- Kutner, M.H., Nachtsheim, C.J., & Neter, J. 2005. *Applied Linear Regression Models*. Fourth Edition. McGraw-Hill Companies, Inc., New York.
- Lawless, J., & Wang, P. 1976. A Simulation Study of Ridge and Other Regression Estimators. *Communications in Statistics-Theory and Methods*, 5(4), 307-323.
- Montgomery, D.C. & E.A. Peck. 1991. *Introduction to Linear Regression Analysis, Second Edition*. New York: John Wiley and Sons, Inc.
- Sembiring, R.K. 2003. *Analisis Regresi*. Edisi Kedua. Bandung: Institut Teknologi Bandung.
- Soemartini. 2008. *Penyelesaian Multikolinearitas Melalui Metode Ridge Regression*. Jawa Barat: UNPAD Jatinangor.
- Sugiyono. 2005. *Metode Penelitian Administrasi*. Bandung: Alfabeta.
- Sumodiningrat. 1994. *Ekonometrika Pengantar*. Yogyakarta: BPF.
- Pratiwi, N.B. 2016. *Perbandingan Regresi Komponen Utama dengan Regresi Ridge untuk Mengatasi Masalah Multikolinearitas*. Skripsi S1. Semarang: Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
- Prenadita, A. 2011. *Penggunaan Metode Ridge Trace dan Variance Inflation Factors (VIF) pada Regresi Ridge*. Skripsi S1. Yogyakarta: Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta.
- Puri, V.G. 2014. *Aplikasi Generalized Ridge Regression untuk Menangani Masalah Multikolinearitas*. Yogyakarta: Jurusan Matematika FMIPA UGM.
- Quadratullah, M.F. 2013. *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi ds2wengan SPSS*. Yogyakarta: ANDI.

LAMPIRAN

Lampiran 1. Data IPM di Provinsi Jawa Tengah Tahun 2016

| IPM | AHH | HLS | RLS | RMG | APS | PDRB | TPAK | KP |
|------------|------------|------------|------------|------------|------------|-------------|-------------|-----------|
| 68.60 | 73.11 | 12.29 | 6.90 | 15.42 | 97.55 | 9677.00 | 66.22 | 796.53 |
| 70.49 | 73.23 | 12.58 | 7.39 | 17.21 | 96.61 | 10554.00 | 65.19 | 1243.32 |
| 67.48 | 72.86 | 11.93 | 6.86 | 17.72 | 92.82 | 9159.00 | 71.68 | 1166.99 |
| 65.52 | 73.69 | 11.40 | 6.26 | 17.58 | 90.29 | 8400.00 | 70.95 | 848.25 |
| 67.41 | 72.87 | 12.61 | 7.05 | 12.21 | 97.13 | 8276.00 | 66.84 | 926.61 |
| 70.66 | 74.14 | 13.05 | 7.66 | 17.21 | 97 | 9497.00 | 64.48 | 688.71 |
| 66.19 | 71.16 | 11.67 | 6.12 | 12.70 | 94.37 | 9877.00 | 72.37 | 792.94 |
| 67.85 | 73.33 | 12.15 | 7.40 | 14.94 | 95.65 | 8501.00 | 74.49 | 1157.86 |
| 72.18 | 75.67 | 12.14 | 7.17 | 13.22 | 97.07 | 12192.00 | 69.96 | 954.93 |
| 73.97 | 76.59 | 12.85 | 8.22 | 13.02 | 99.49 | 11227.00 | 66.93 | 1774.39 |
| 75.06 | 77.46 | 13.79 | 8.58 | 13.25 | 99.35 | 10452.00 | 67.29 | 1867.31 |
| 68.23 | 75.88 | 12.43 | 6.57 | 8.96 | 97.99 | 8589.00 | 71.22 | 522.38 |
| 74.90 | 77.11 | 13.64 | 8.49 | 12.43 | 97.69 | 10722.00 | 70.24 | 1118.91 |
| 71.43 | 75.43 | 12.30 | 6.87 | 12.59 | 94.71 | 11688.00 | 71.12 | 931.96 |
| 68.52 | 74.37 | 12.26 | 6.62 | 14.57 | 94.26 | 9487.00 | 72.15 | 687.5 |
| 66.61 | 73.88 | 11.92 | 6.18 | 12.99 | 99.44 | 8846.00 | 70.21 | 478.8 |
| 68.60 | 74.27 | 12.03 | 6.93 | 13.62 | 100 | 9453.00 | 70.78 | 615.42 |
| 69.03 | 75.69 | 11.92 | 6.83 | 11.85 | 95.64 | 9548.00 | 66.83 | 831.54 |
| 72.94 | 76.43 | 13.19 | 7.85 | 14.57 | 97.26 | 10348.00 | 71.75 | 1979.21 |
| 70.25 | 75.67 | 12.28 | 7.32 | 14.02 | 92.4 | 9695.00 | 69.85 | 1200.8 |
| 70.10 | 75.27 | 12.44 | 7.46 | 12.29 | 92.45 | 9377.00 | 67.73 | 1258.37 |
| 72.40 | 75.54 | 12.83 | 7.48 | 14.39 | 95.65 | 11102.00 | 76.37 | 1071.12 |
| 67.60 | 75.39 | 12.06 | 6.55 | 13.85 | 96.01 | 8593.00 | 74.37 | 864.7 |
| 70.11 | 74.20 | 12.68 | 6.65 | 15.34 | 96.04 | 10631.00 | 66.49 | 947.53 |
| 66.38 | 74.46 | 11.51 | 6.42 | 16.11 | 95.63 | 8568.00 | 67.7 | 950.28 |
| 67.71 | 73.41 | 12.15 | 6.56 | 16.02 | 87.9 | 9300.00 | 70.98 | 1052.58 |
| 64.17 | 72.87 | 11.87 | 6.05 | 15.76 | 89.61 | 7447.00 | 65.57 | 1277.41 |
| 65.84 | 71.02 | 12.01 | 6.54 | 16.00 | 91.95 | 8709.00 | 66.41 | 1624.86 |
| 63.98 | 68.41 | 11.37 | 6.17 | 15.67 | 95.79 | 9148.00 | 67.42 | 1079.11 |
| 77.16 | 76.62 | 13.55 | 10.29 | 15.38 | 98.37 | 11090.00 | 65.32 | 6683.89 |
| 80.76 | 77.03 | 14.50 | 10.37 | 13.60 | 99.22 | 13900.00 | 66.1 | 11677.74 |
| 81.14 | 76.87 | 14.98 | 9.82 | 36.02 | 100 | 14811.00 | 70.53 | 3520.02 |
| 81.19 | 77.21 | 14.70 | 10.49 | 5.47 | 98.59 | 13909.00 | 69.87 | 4627.3 |
| 73.32 | 74.15 | 12.77 | 8.29 | 14.38 | 96.04 | 11721.00 | 69.28 | 6655.29 |
| 73.55 | 74.18 | 12.88 | 8.28 | 24.10 | 93.21 | 11849.00 | 66.33 | 7167.64 |

Lampiran 2. Data Transformasi

| IPM* | AHH* | HLS* | RLS* | RMG* | APS* | PDRB* | TPAK* | KP* |
|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|------------|
| -0.449 | -0.742 | -0.345 | -0.450 | 0.094 | 0.586 | -0.300 | -1.016 | -0.512 |
| -0.027 | -0.681 | -0.014 | -0.048 | 0.482 | 0.270 | 0.222 | -1.371 | -0.327 |
| -0.700 | -0.871 | -0.757 | -0.483 | 0.593 | -1.003 | -0.608 | 0.863 | -0.358 |
| -1.138 | -0.445 | -1.362 | -0.976 | 0.563 | -1.853 | -1.060 | 0.612 | -0.491 |
| -0.715 | -0.866 | 0.020 | -0.327 | -0.601 | 0.445 | -1.134 | -0.803 | -0.458 |
| 0.011 | -0.214 | 0.523 | 0.174 | 0.482 | 0.401 | -0.407 | -1.615 | -0.557 |
| -0.988 | -1.743 | -1.054 | -1.091 | -0.495 | -0.482 | -0.181 | 1.101 | -0.514 |
| -0.617 | -0.630 | -0.505 | -0.040 | -0.010 | -0.052 | -1.000 | 1.831 | -0.362 |
| 0.351 | 0.572 | -0.517 | -0.229 | -0.383 | 0.425 | 1.197 | 0.271 | -0.447 |
| 0.751 | 1.044 | 0.294 | 0.634 | -0.426 | 1.238 | 0.622 | -0.772 | -0.106 |
| 0.995 | 1.491 | 1.368 | 0.929 | -0.376 | 1.191 | 0.161 | -0.648 | -0.067 |
| -0.532 | 0.680 | -0.185 | -0.721 | -1.306 | 0.734 | -0.947 | 0.705 | -0.626 |
| 0.959 | 1.311 | 1.197 | 0.855 | -0.554 | 0.633 | 0.322 | 0.368 | -0.378 |
| 0.183 | 0.449 | -0.334 | -0.475 | -0.519 | -0.368 | 0.897 | 0.671 | -0.456 |
| -0.467 | -0.096 | -0.380 | -0.680 | -0.090 | -0.519 | -0.413 | 1.025 | -0.558 |
| -0.894 | -0.347 | -0.768 | -1.042 | -0.432 | 1.221 | -0.795 | 0.357 | -0.644 |
| -0.449 | -0.147 | -0.642 | -0.426 | -0.296 | 1.409 | -0.433 | 0.554 | -0.588 |
| -0.353 | 0.582 | -0.768 | -0.508 | -0.680 | -0.055 | -0.377 | -0.806 | -0.498 |
| 0.521 | 0.962 | 0.683 | 0.330 | -0.090 | 0.489 | 0.099 | 0.887 | -0.021 |
| -0.080 | 0.572 | -0.357 | -0.105 | -0.209 | -1.144 | -0.289 | 0.233 | -0.344 |
| -0.114 | 0.366 | -0.174 | 0.010 | -0.584 | -1.127 | -0.479 | -0.496 | -0.320 |
| 0.400 | 0.505 | 0.272 | 0.026 | -0.129 | -0.052 | 0.548 | 2.478 | -0.398 |
| -0.673 | 0.428 | -0.608 | -0.738 | -0.246 | 0.069 | -0.945 | 1.789 | -0.484 |
| -0.112 | -0.183 | 0.100 | -0.656 | 0.077 | 0.079 | 0.268 | -0.923 | -0.450 |
| -0.946 | -0.049 | -1.236 | -0.844 | 0.244 | -0.059 | -0.960 | -0.507 | -0.448 |
| -0.648 | -0.588 | -0.505 | -0.729 | 0.224 | -2.656 | -0.524 | 0.622 | -0.406 |
| -1.440 | -0.866 | -0.825 | -1.148 | 0.168 | -2.082 | -1.627 | -1.240 | -0.313 |
| -1.067 | -1.815 | -0.665 | -0.746 | 0.220 | -1.295 | -0.876 | -0.951 | -0.168 |
| -1.482 | -3.155 | -1.396 | -1.050 | 0.149 | -0.005 | -0.615 | -0.603 | -0.395 |
| 1.465 | 1.059 | 1.094 | 2.334 | 0.086 | 0.862 | 0.541 | -1.326 | 1.934 |
| 2.270 | 1.270 | 2.180 | 2.399 | -0.300 | 1.147 | 2.213 | -1.058 | 4.008 |
| 2.355 | 1.188 | 2.728 | 1.948 | 4.560 | 1.409 | 2.755 | 0.467 | 0.619 |
| 2.366 | 1.362 | 2.408 | 2.498 | -2.063 | 0.936 | 2.218 | 0.240 | 1.079 |
| 0.606 | -0.209 | 0.203 | 0.691 | -0.131 | 0.079 | 0.916 | 0.037 | 1.922 |
| 0.658 | -0.193 | 0.329 | 0.683 | 1.976 | -0.872 | 0.992 | -0.978 | 2.135 |

Lampiran 3. *Output* Program R

```

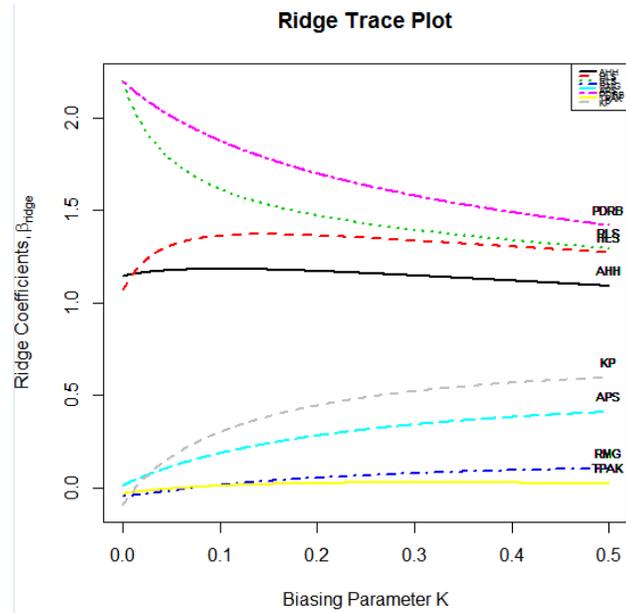
Residuals:
      Min       1Q   Median       3Q      Max
-0.43923 -0.10673  0.03186  0.12895  0.26423

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.059e+00  2.142e+00   2.829  0.00888 **
AHH          4.442e-01  2.693e-02  16.499  2.72e-15 ***
HLS          9.230e-01  1.149e-01   8.034  1.63e-08 ***
RLS          1.365e+00  9.723e-02  14.041  1.20e-13 ***
RMG          -7.729e-03  8.240e-03  -0.938  0.35688
APS          3.367e-03  1.438e-02   0.234  0.81670
PDRB         9.873e-04  3.775e-05  26.154 < 2e-16 ***
TPAK         -8.111e-03  1.291e-02  -0.628  0.53523
KP           -2.880e-05  2.565e-05  -1.123  0.27177
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2001 on 26 degrees of freedom
Multiple R-squared:  0.9985,    Adjusted R-squared:  0.9981
F-statistic: 2181 on 8 and 26 DF,  p-value: < 2.2e-16

```

| | k values |
|------------------------------|----------|
| Thisted (1976): | 0.00096 |
| Dwividi & Srivastava (1978): | 0.00016 |
| LW (lm.ridge) | 0.01159 |
| LW (1976) | 0.00044 |
| HKB (1975) | 0.00127 |
| Kibria (2003) (AM) | 1.75781 |
| Minimum GCV at | 0.00200 |
| Minimum CV at | 0.04200 |
| Kibria 2003 (GM): | 0.00648 |
| Kibria 2003 (MED): | 0.00890 |
| Muniz et al. 2009 (KM2): | 65.38815 |
| Muniz et al. 2009 (KM3): | 0.80688 |
| Muniz et al. 2009 (KM4): | 12.42116 |
| Muniz et al. 2009 (KM5): | 0.08051 |
| Muniz et al. 2009 (KM6): | 10.73587 |
| Mansson et al. 2012 (KMN8): | 65.43844 |
| Mansson et al. 2012 (KMN9): | 0.35106 |
| Mansson et al. 2012 (KMN10): | 14.04037 |
| Mansson et al. 2012 (KMN11): | 0.07122 |
| Mansson et al. 2012 (KMN12): | 11.04039 |
| Dorugade et al. 2010: | 0.00000 |
| Dorugade et al. 2014: | 0.00000 |



Call:

```
lmridge.default(formula = IPM ~ ., data = data, K = c(0.03471074))
```

Coefficients: for Ridge parameter K= 0.03471074

| | Estimate | Estimate (Sc) | StdErr (Sc) | t-value (Sc) | Pr(> t) |
|-----------|------------|---------------|-------------|--------------|--------------|
| Intercept | -1.621e-10 | 4.454e-10 | 6.007e+00 | 0.000 | 1.89e-09 *** |
| AHH | 1.978e-01 | 1.170e+00 | 7.177e-02 | 16.302 | < 2e-16 *** |
| HLS | 2.153e-01 | 1.274e+00 | 1.041e-01 | 12.237 | < 2e-16 *** |
| RLS | 3.128e-01 | 1.851e+00 | 1.135e-01 | 16.303 | < 2e-16 *** |
| RMG | -4.260e-03 | -2.520e-02 | 5.464e-02 | -0.461 | 0.645 |
| APS | 1.498e-02 | 8.861e-02 | 6.115e-02 | 1.449 | 0.147 |
| PDRB | 3.475e-01 | 2.056e+00 | 8.424e-02 | 24.403 | < 2e-16 *** |
| TPAK | -1.944e-03 | -1.150e-02 | 5.433e-02 | -0.212 | 0.832 |
| KP | 1.881e-02 | 1.113e-01 | 7.605e-02 | 1.463 | 0.143 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ridge Summary

| R2 | adj-R2 | DF ridge | F | AIC | BIC |
|---------|---------|----------|------------|------------|-----------|
| 0.97680 | 0.97070 | 7.12806 | 1620.70728 | -201.62306 | -66.09927 |

Ridge minimum MSE= 0.2825796 at K= 0.03471074

Call:

```
lmridge.default(formula = IPM ~ ., data = data, K = c(0.01203737))
```

Coefficients: for Ridge parameter K= 0.01203737

| | Estimate | Estimate (Sc) | StdErr (Sc) | t-value (Sc) | Pr(> t) | |
|-----------|------------|---------------|-------------|--------------|----------|-----|
| Intercept | -2.057e-10 | 1.874e-10 | 6.007e+00 | 0.000 | 1.89e-09 | *** |
| AHH | 1.954e-01 | 1.156e+00 | 6.753e-02 | 17.122 | < 2e-16 | *** |
| HLS | 1.978e-01 | 1.170e+00 | 1.146e-01 | 10.213 | < 2e-16 | *** |
| RLS | 3.443e-01 | 2.037e+00 | 1.310e-01 | 15.544 | < 2e-16 | *** |
| RMG | -6.817e-03 | -4.033e-02 | 4.980e-02 | -0.810 | 0.418 | |
| APS | 7.137e-03 | 4.222e-02 | 5.601e-02 | 0.754 | 0.451 | |
| PDRB | 3.620e-01 | 2.141e+00 | 8.082e-02 | 26.497 | < 2e-16 | *** |
| TPAK | -3.964e-03 | -2.345e-02 | 4.933e-02 | -0.475 | 0.635 | |
| KP | -7.824e-04 | -4.629e-03 | 7.549e-02 | -0.061 | 0.951 | |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Ridge Summary

| R2 | adj-R2 | DF ridge | F | AIC | BIC |
|---------|---------|----------|------------|------------|-----------|
| 0.99050 | 0.98810 | 7.63636 | 2115.06930 | -210.14032 | -73.82594 |

Ridge minimum MSE= 0.103325 at K= 0.01203737

```
> mod3 <- lmridge(IPM~., data=data, K = c(0.03471074) )  
> rstats1(mod3)
```

Ridge Regression Statistics 1:

| | Variance | Bias^2 | MSE | rsigma2 | F | R2 | adj-R2 | CN |
|--------------|----------|--------|--------|---------|----------|--------|--------|---------|
| K=0.03471074 | 0.0514 | 0.2312 | 0.2826 | 0.0027 | 1620.707 | 0.9768 | 0.9707 | 47.5639 |

```
> mod4 <- lmridge(IPM~., data=data, K = c(0.01203737) )  
> rstats1(mod4)
```

Ridge Regression Statistics 1:

| | Variance | Bias^2 | MSE | rsigma2 | F | R2 | adj-R2 | CN |
|--------------|----------|--------|--------|---------|----------|--------|--------|---------|
| K=0.01203737 | 0.0551 | 0.0482 | 0.1033 | 0.0021 | 2115.069 | 0.9905 | 0.9881 | 63.9704 |

```
> mod6 <- lmridge(IPM~., data=data, scaling="sc", K=c(0.03471074) )  
> vif(mod6)
```

| | AHH | HLS | RLS | RMG | APS | PDRB | TPAK |
|--------------|----------|----------|----------|----------|----------|----------|---------|
| k=0.03471074 | 1.910991 | 4.019174 | 4.780993 | 1.107549 | 1.387437 | 2.633035 | 1.09523 |
| | KP | | | | | | |
| k=0.03471074 | 2.145469 | | | | | | |

```
> mod7 <- lmridge(IPM~., data=data, scaling="sc", K=c(0.01203737) )  
> vif(mod7)
```

| | AHH | HLS | RLS | RMG | APS | PDRB | TPAK |
|--------------|----------|----------|---------|----------|----------|----------|----------|
| k=0.01203737 | 2.20808 | 6.357016 | 8.31416 | 1.200966 | 1.519078 | 3.162457 | 1.178016 |
| | KP | | | | | | |
| k=0.01203737 | 2.758946 | | | | | | |

```

> mod<-lmridge(IPM~., data=data, K=c(0.03471074), scaling="sc")
> ## Scaled Coefficients
> mod$coef
      K=0.03471074
AHH    1.16998342
HLS    1.27361831
RLS    1.85074498
RMG   -0.02520362
APS    0.08860856
PDRB   2.05582870
TPAK   -0.01150329
KP     0.11126214
> ## Re-Scaled Coefficients
> coef(mod)
      Intercept          AHH          HLS          RLS          RMG
-1.621362e-10  1.977633e-01  2.152808e-01  3.128330e-01 -4.260189e-03
              APS          PDRB          TPAK          KP
 1.497758e-02  3.474985e-01 -1.944411e-03  1.880673e-02
. |
> mod<-lmridge(IPM~., data=data, K=c(0.01203737), scaling="sc")
> ## Scaled Coefficients
> mod$coef
      K=0.01203737
AHH    1.15627353
HLS    1.17024967
RLS    2.03688026
RMG   -0.04033228
APS    0.04222384
PDRB   2.14146981
TPAK   -0.02344904
KP     -0.00462853
> ## Re-Scaled Coefficients
> coef(mod)
      Intercept          AHH          HLS          RLS          RMG
-2.057362e-10  1.954459e-01  1.978083e-01  3.442956e-01 -6.817400e-03
              APS          PDRB          TPAK          KP
 7.137132e-03  3.619745e-01 -3.963611e-03 -7.823644e-04
> x$kH
[1] 0.03471074
> x$kLW
[1] 0.01203737

```

