



KEMENTERIAN
KESEHATAN
REPUBLIK
INDONESIA



Analisis Data Menggunakan Multiple Linear Regression Test di Bidang Kesehatan Masyarakat dan Klinis

Dr. Rr. Nur Fauziah, SKM, MKM, RD

**Penerbit
Politeknik Kesehatan Kemenkes Bandung**

ISBN 978-623-91302-6-8



9 786239 130268

Analisis Data Menggunakan Multiple Linear Regression Test di Bidang Kesehatan Masyarakat dan Klinis

Dr. Rr. Nur Fauziyah, SKM, MKM, RD

**Penerbit
Politeknik Kesehatan Kemenkes Bandung**

Analisis Data Menggunakan Multiple Linear Regression Test di Bidang Kesehatan Masyarakat dan Klinis

Penulis :

Dr. Rr. Nur Fauziah, SKM, MKM, RD

ISBN : 978-623-91302-6-8

Editor :

Gurid Pramintarto Eko Mulyo, SKM, M.Sc

Penyunting :

Surmita, S.Gz, M.Kes

Desain sampul dan Tata Letak :

Azimah Istianah, S.Ds

Penerbit :

Politeknik Kesehatan Kemenkes Bandung

Redaksi :

Jln. Pajajaran No 56

Bandung 40171

Tel (022) 4231627

Fax (022) 4231640

Email : info@poltekkesbandung.ac.id

Cetakan pertama, Maret 2020

Hak cipta dilindungi undang-undang

Dilarang diperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa izin tertulis dari penerbit

KATA PENGANTAR

Buku pengolahan dan analisis data telah banyak tersedia, namun hanya sedikit yang memberikan contoh-contoh nyata bidang kesehatan dan kedokteran yang mudah dipahami oleh peneliti dan mahasiswa bidang kesehatan. Buku ini yang berjudul “Analisis Data Menggunakan Multiple Linear Regression Test di Bidang Kesehatan Masyarakat dan Klinis”.

Buku ini disusun secara sistematis dan rinci disertai contoh nyata di bidang kesehatan masyarakat dan klinis, yang dipandu selangkah demi selangkah dalam tahap- tahap penyelesaiannya. Pada bagian akhir analisis, diberikan contoh bagaimana cara penyajian data dalam bentuk tabel dan bagaimana menuliskan interpretasinya.

Semoga buku ini bermanfaat bagi peneliti dan mahasiswa bidang kesehatan dan kedokteran untuk membantu dalam pengolahan dan analisa data, skripsi, thesis, disertasi maupun analisa data untuk monitoring dan evaluasi program kesehatan. Kritik dan saran kami terima dengan senang hati untuk kesempurnaan buku ini.

Bandung, Maret 2020

Penulis,

Dr. Rr. Nur Fauziah, SKM, MKM, RD

DAFTAR ISI

	Halaman
KATA PENGANTAR	1
DAFTAR ISI	2
1. Konsep Analisis Multivariat	3
2. Konfounding dan Interaksi	5
3. Konsep Analisis Regresi Linier Ganda	6
3.1. Asumsi Analisis Regresi Linier Ganda	7
3.2. Kegunaan Analisis Regresi Linier Ganda	8
4. Pemodelan Regresi Linier Multivariat	9
5. Uji Diagnostik Asumsi Regresi Linier Multivariat	12
6. Aplikasi Pemodelan Regresi Linier Multivariat	13
6.1. Seleksi Bivariat	14
6.2. Pemodelan Regresi Linier Multivariat	19
6.3. Uji Konfounding	22
6.4. Uji Asumsi Regresi Linier Multivariat	26
6.5. Uji Interaksi Regresi Linier	34
6.6. Model Akhir dan Interpretasi	35
DAFTAR PUSTAKA	65

1. Konsep Analisis Multivariat

Setelah kita mempelajari dan menguasai teknik analisis bivariat maka teknik analisis berikutnya yang perlu diketahui adalah analisis multivariat. Analisis multivariat merupakan teknik analisis perluasan atau pengembangan dari analisis bivariat. Kalau analisis bivariat melihat hubungan atau keterkaitan dua variabel, maka teknik analisis multivariat bertujuan melihat mempelajari hubungan beberapa variabel (lebih dari satu variabel) independen dengan satu atau beberapa variabel dependen (umumnya satu variabel dependen).

Setelah mempelajari bagian ini, peserta mampu mengetahui 1) Konsep analisis multivariat pada data numerik, 2) Aplikasi dan kegunaan analisis regresi linier multivariat, 3) Aplikasi dan Cara uji confounding, 4) Aplikasi dan Cara uji interaksi, 5) Aplikasi Regresi Linier Ganda, 6) Aplikasi Uji Asumsi Regresi Linier Ganda.

Apa pentingnya mempelajari multivariat?, seperti kita ketahui bersama bahwa dalam bidang apapun, termasuk dalam hal ini bidang kesehatan masyarakat, suatu akibat/fenomena tertentu tidak mungkin disebabkan oleh hanya satu penyebab, kenyataan yang ada memperlihatkan bahwa hampir semua kejadian atau fenomena yang ada di dunia ini disebabkan oleh banyak penyebab (beberapa faktor atau multi faktor).

Sebagai contoh fenomena atau kejadian kelahiran bayi BBLR (Berat Bayi Lahir Rendah), tidak mungkin kejadian bayi BBLR hanya dipengaruhi satu faktor saja misalnya berat ibu atau status gizi ibu, anemia, penyakit lain yang diderita ibu, atau sosial ekonomi rumah tangga. Kita ketahui bahwa faktor penyebab bayi BBLR tidak hanya berat ibu tapi bisa jadi karena faktor gizi ibu, riwayat menderita hipertensi, pemeriksaan ANC, dan lain-lain. Karena banyak dan kompleksnya faktor independen yang mempengaruhi variabel dependen maka sangatlah penting mempelajari dan menguasai teknik analisis multivariat.

Proses analisis multivariat dilakukan dengan menghubungkan beberapa variabel independen dengan satu variabel dependen pada waktu yang bersamaan. Jumlah sampel dalam analisis multivariat sangat penting diperhatikan, sebaiknya jangan terlalu sedikit, pedoman yang berlaku umum adalah setiap variabel independen minimal diperlukan 10 sampai 20 responden. Bila dalam model multivariate terdapat 10 variabel independen, maka diperlukan jumlah sampel minimal = 10×10 responden = 100 responden.

Referensi lain menyebutkan untuk satu variabel independen minimal diperlukan 10 sampai 20 kejadian kasus (outcome=1) dan minimal diperlukan 10 sampai 20 kontrol (outcome=0) sebagai pembanding. Bila dalam model multivariate terdapat 10 variabel independen, maka diperlukan jumlah

sampel minimal = 10 x 10 kasus = 100 kasus dan 10 x 10 kontrol = 100 kontrol, total 200 kasus dan kontrol.

Dengan melakukan analisis multivariat kita dapat mengetahui hal-hal berikut:

- a. Variabel independent apa saja yang betul-betul berhubungan dengan variable dependen? (bukan hanya hubungan semu seperti pada analisis bivariat)
- b. Variabel independen mana yang paling besar pengaruhnya terhadap variabel dependen?
- c. Apakah hubungan suatu variabel independen dengan variabel dependen dipengaruhi oleh variabel lain atau tidak? (ada konfounding atau tidak)
- d. Apakah hubungan suatu variabel independen dengan variabel dependen berbeda menurut kategori variabel lain atau tidak? (ada interaksi atau tidak)
- e. Bagaimana bentuk hubungan variabel independen dengan variabel dependen, apakah berpengaruh langsung atau pengaruh tidak langsung ?.

Prosedur pengujian tergantung dari jenis data yang diuji apakah katagori atau numerik. Berikut adalah gambaran secara garis besar beberapa analisis statistik yang dapat digunakan untuk analisis multivariat:

Variabel Independen	Variabel Dependen	Jenis Uji
Numerik & Kategorik (variabel numeric lebih banyak)	Numerik	Uji Regresi Linier
Numerik & Kategorik	Kategorik	Uji Regresi Logistik
Numerik & Kategorik	Numerik waktu	Uji Regresi Cox
Numerik	Kategorik	Uji Diskriminan
Kategorik	Numerik	MANOVA

Perbedaan antara regresi linear dengan regresi logistik terletak pada jenis variabel dependennya. Regresi linear digunakan apabila variabel dependennya numerik, sedangkan regresi logistik digunakan pada data yang dependennya kategorik. Persamaan antara regresi linier dan logistic terletak pada variable independennya, variable independennya merupakan gabungan antara data numeric dan kategorik. Namun pada regresi linier harus diupayakan variabel independen numerik lebih banyak dibanding kategorik agar lebih mudah dalam memenuhi asumsi pemodelannya.

2. Konfounding dan Interaksi

Sebelum melakukan analisis multivariat kita harus mengetahui terlebih dahulu mengenai konsep konfounding dan Interaksi.

Konfounding

Konfounding merupakan kondisi bias dalam mengestimasi efek pajanan atau exposure terhadap kejadian penyakit atau masalah kesehatan, akibat dari perbandingan yang tidak seimbang kelompok expose dengan kelompok non expose pada variable perancu atau variable ketiga. Masalah ini terjadi karena pada dasarnya sudah ada perbedaan risiko terjadinya penyakit pada variable perancu, risiko terjadinya penyakit pada kedua kelompok di variable perancu itu berbeda, meskipun expose dihilangkan.

Suatu variabel dikatakan konfounding jika variabel tersebut merupakan faktor risiko untuk terjadinya penyakit (outcome) dan berhubungan dengan variabel independen tapi tidak merupakan hasil dari variabel independent (tidak variable antara).

Interaksi

Interaksi atau sering disebut juga efek modifikasi adalah heterogenitas efek dari satu exposure menurut kategori variable lainnya atau pada tingkat expose yang lain. Jadi efek satu exposure pada kejadian penyakit berbeda pada kelompok variabel lainnya. Tidak adanya modifikasi efek, berarti efek exposure homogen. Modifikasi efek merupakan konsep yang penting dalam analisis data karena pada saat analisis kita harus menentukan apakah akan melaporkan efek gabungan atau efek yang terpisah untuk masing-masing strata. Apabila ada interaksi maka yang dilaporkan adalah efek setelah berinteraksi.

Pada analisis multivariat, jika ditemukan adanya interaksi antar variabel expose dengan variabel lainnya, maka nilai koefisien B, harus dilaporkan setelah variabel tersebut berinteraksi.

Bila tidak ada interaksi, maka persamaan regresi linier adalah sebagai berikut:

Persamaan regresi linier multivariat:

$$(Y) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_nX_n$$

Keterangan:

Y = Variabel dependen (numerik)

β_0 = intercept

β_i = koefisien regresi variabel X_i

X_i = Variabel independen (numerik atau kategorik)

$$(Y) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_1 \cdot X_2 + \dots + \beta_nX_n$$

Jika ada interaksi, maka persamaan regresi linier adalah sebagai berikut: Persamaan regresi linier multivariat:

$$(Y) = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_1 \cdot X_2 + \dots + \beta_nX_n$$

Keterangan:

Y = Variabel dependen (numerik)

β_0 = intercept

β_i = koefisien regresi variabel X_1, X_2, \dots, X_n

X_1 = Variabel independen-1 (numerik atau kategorik)

X_2 = Variabel independen-2 (numerik atau kategorik)

$X_1 \cdot X_2$ = Perkalian atau Interaksi variabel independen-1 (numerik atau kategorik)
dengan variabel independen-2 (numerik atau kategorik)

3. Konsep Analisis Regresi Linier Ganda

Analisis *Multiple regression Linear* atau sering disebut juga analisis regresi linier ganda merupakan perluasan analisis *Simple Linear Regression* (regresi linier sederhana). Dalam analisis *Simple Linear Regression* hanya ada satu variabel independen (variabel bebas) dihubungkan dengan satu variabel dependen (terikat).. Sedangkan pada *Multiple regression Linear* merupakan analisis hubungan antara beberapa variabel independen dengan satu variabel dependen. Misalkan untuk mengetahui

faktor-faktor yang berhubungan dengan tekanan darah, dilakukan analisis dengan melibatkan variabel independen: umur, berat badan, dan jenis kelamin.

Dalam regresi linier ganda variabel dependennya harus numerik sedangkan variabel independen boleh semuanya numerik dan boleh juga campuran numerik dan katagorik. Model persamaan regresi linier ganda merupakan perluasan regresi linier sederhana, yaitu:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k + e$$

3.1. ASUMSI REGRESI LINIER GANDA

Seperti pada umumnya pengujian statistik, dari analisis regresi linier ganda diharapkan dapat memberikan informasi yang lebih banyak bukan sekedar diskripsi data teramati. Kita tentu ingin menarik inferensi (menggeneralisasi) tentang hubungan variabel-variabel dalam populasi asal dari sampel diambil. Bagaimanakah hubungan antara umur, berat badan dan jenis kelamin 'pada semua orang (populasi)', tidak hanya seperti yang teramati di sejumlah orang pada sampel?. Oleh karena itu agar inferensi kita valid maka dalam analisis regresi dianjurkan untuk mengikuti kaidah-kaidah yang dipersyaratkan dalam analisis regresi. Dengan kata lain, setiap melakukan analisis *Multiple regression Linear* harus memenuhi asumsi/persyaratan yang ditetapkan. Adapun asumsi yang digunakan dalam *Multiple regression Linear* sebagai berikut

a. Asumsi Eksistensi (Variabel Random)

Untuk tiap nilai dari variabel X (variabel independen), variabel Y (dependen) adalah variabel random yang mempunyai mean dan varian tertentu. Asumsi ini berkaitan dengan teknik pengambilan sampel. Untuk memenuhi asumsi ini, sampel yang diambil harus dilakukan secara random. Cara mengetahui asumsi eksistensi dengan cara melakukan analisis deskriptif variabel residual dari model, bila residual menunjukkan adanya mean dan sebaran (varian atau standar deviasi) maka asumsi eksistensi terpenuhi.

b. Asumsi Independensi

Suatu keadaan dimana masing-masing nilai Y bebas satu sama lain. Jadi nilai dari tiap-tiap individu saling berdiri sendiri. Tidak diperbolehkan nilai observasi yang berbeda yang diukur dari satu individu diukur dua kali. Untuk mengetahui asumsi ini dilakukan dengan cara mengeluarkan uji Durbin Watson, bila nilai Durbin -2 s.d. $+2$ berarti asumsi independensi terpenuhi, sebaliknya bila nilai Durbin < -2 atau $> +2$ berarti asumsi tidak terpenuhi.

c. Asumsi Linieritas

Nilai mean dari variabel Y untuk suatu kombinasi $X_1, X_2, X_3, \dots, X_k$ terletak pada garis/bidang linier yang dibentuk dari persamaan regresi. Untuk mengetahui asumsi linieritas dapat diketahui dari uji ANOVA (overall F test) bila hasilnya signifikan ($p \text{ value} < \alpha$) maka model berbentuk linier.

d. Asumsi Homoscedasticity

Varian nilai variabel Y sama untuk semua nilai variabel X. Homoscedasticity dapat diketahui dengan melakukan pembuatan plot residual. Bila titik tebaran tidak berpola tertentu dan menyebar merata disekitar garis titik nol maka dapat disebut varian homogen pada setiap nilai X dengan demikian asumsi homoscedasticity terpenuhi. Sebaliknya bila titik tebaran membentuk pola tertentu misalnya mengelompok di bawah atau di atas garis tengah nol, maka diduga variannya terjadi heteroscedasticity.

e. Asumsi Normalitas

Variabel Y mempunyai distribusi normal untuk setiap pengamatan variabel X. dapat diketahui dari Normal P-P Plot residual, bila data menyebar di sekitar garis diagonal dan mengikuti arah garis diagonal, maka model regresi memenuhi asumsi model regresi tidak memenuhi asumsi normalitas.

3.2. KEGUNAAN ANALISIS REGRESI LINIER GANDA

Tujuan analisis regresi linier ganda adalah untuk menemukan model regresi yang paling sesuai untuk mengetahui faktor-faktor yang berhubungan dengan variabel dependen. Pada prinsipnya, model regresi ganda dapat berguna untuk dua hal:

1. **Model Prediksi**, Model regresi yang berisi variabel independen apa saja yang signifikan dapat digunakan untuk memprediksi variabel dependen numerik. Dalam hal ini dapat diestimasi atau diprediksi nilai variabel dependen bila seseorang/individu mempunyai suatu set variabel independen tertentu. Misalnya kita melakukan analisis variabel independen umur, BB dan jenis kelamin dihubungkan dengan variabel dependen tekanan darah. Dari hasil regresi, seseorang individu dapat diperkirakan tekanan darahnya pada umur, berat badan, dan jenis kelamin tertentu. Dari model prediksi ini dapat diketahui variabel mana yang paling besar atau dominan mempengaruhi variabel dependen, yang ditunjukkan dari nilai koefisien regresi (B) yang sudah distandardisasi yaitu nilai **beta** yang terbesar.
2. **Model Asosiasi**, Model regresi untuk mengkuantifikasi asosiasi atau hubungan satu atau beberapa variabel independen dengan satu variabel dependen numerik. Pada fungsi ini regresi linier dapat

digunakan untuk mengetahui variabel independen apa saja yang berhubungan dengan variabel dependen. Selain itu kita juga dapat mengetahui seberapa besar hubungan masing-masing independen terhadap variabel independen lainnya.

4. Pemodelan Regresi Linier Multivariat

Satu hal yang penting dalam regresi ganda adalah bagaimana memilih variabel independen sehingga terbentuk sebuah model yang paling sesuai untuk menjelaskan atau menggambarkan variabel dependen yang sesungguhnya di populasi.

Dalam pembuatan model seringkali dijumpai pandangan yang kurang tepat yaitu "memasukkan semua variabel atau sebanyak mungkin variabel independen ke dalam model". Alasannya, dengan memasukkan sebanyak mungkin variabel independen ke dalam model, maka variabel dependen diharapkan dapat diprediksi dengan sempurna. Perlu diketahui bahwa penambahan variabel independen tidak selalu meningkatkan kemampuan prediksi variabel independen terhadap variabel dependen, sebab semakin banyak variabel independen (lebih-lebih variabel yang tidak relevan) mengakibatkan makin besarnya nilai standar error (Se). disamping itu, model dengan banyak variabel terkadang menyulitkan dalam interpretasi.

Berdasarkan pertimbangan tersebut pemilihan variabel independen hendaknya dengan memperhatikan aspek substansi terlebih dahulu kemudian baru memperhatikan aspek statistik signifikan. Model yang dihasilkan diharapkan model yang PARSIMONI, artinya variabel yang masuk dalam model sebaiknya yang sedikit jumlahnya, namun cukup baik untuk menjelaskan faktor-faktor penting yang berhubungan dengan variabel dependen.

Banyak kriteria yang dapat digunakan untuk memilih variabel masuk dalam model, salah satu kriteria yang sering digunakan adalah melihat perubahan R^2 (*R Square*). Namun penggunaan kriteria ini perlu hati-hati, karena setiap penambahan satu variabel independen akan meningkatkan R^2 walaupun variabel tersebut tidak cukup penting. Oleh karena itu model yang digunakan adalah model dengan nilai R^2 yang besar namun variabel independen jumlahnya sedikit.

Berikut langkah-langkah dalam pemodelan regresi linier ganda:

- 1). SELEKSI BIVARIAT: Melakukan analisis bivariat untuk menentukan variabel yang menjadi kandidat model multivariat. Masing-masing variabel independen dihubungkan dengan variabel dependen (bivariat), bila hasil uji bivariat mempunyai nilai $p < 0,25$, maka variabel

tersebut masuk dalam model multivariat. Untuk variabel yang p value-nya $> 0,25$ namun secara substansi penting, maka variabel tersebut dapat masuk ke kandidat model multivariat.

METODE PEMODELAN TERBARU MENYARANKAN UNTUK TIDAK MELAKUKAN SELEKSI BIVARIAT. Seleksi bivariat hanya perlu kalau tidak tersedia program komputer yang mampu melakukan perhitungan dengan cepat dalam hitungan detik. Pada masa lalu program komputer untuk analisis statistik harus diketik secara manual dan di jalankan (run) kemudian ditunggu berjam jam (bahkan menunggu sampai besok pagi) untuk mendapatkan outputnya. Belum tentu sekali jalan berhasil mendapatkan output, bisa jadi error dan harus diulang kembali sampai berkali-kali.

- 2) PEMODELAN: Dimulai dengan FULL MODEL dengan memasukkan semua variabel yang lolos seleksi bivariat (atau memasukkan semua variabel independen). Kemudian REDUCE MODEL, yakni memilih variabel yang signifikan untuk masuk dalam model, dengan cara mengeluarkan variabel yang sig. $> 0,05$ dan mempertahankan variabel yang mempunyai sig. $< 0,05$. Pengeluaran variabel dilakukan secara bertahap, satu per satu, tidak serentak, namun dilakukan secara bertahap mulai dari variabel yang mempunyai sig. terbesar. Pada variabel yang kategorinya lebih dari dua (dummy variabel), nilai sig. $< 0,05$ atau sig. $> 0,05$ dilihat dari nilai sig. yang terkecil. Apabila pengeluaran variabel yang sig.nya $> 0,05$ menyebabkan perubahan nilai **koefisien B** pada variabel lainnya lebih dari 10%, maka variabel tadi harus dimasukkan ke dalam model kembali, tidak boleh dikeluarkan, karena variabel tadi merupakan konfounding (IDENTIFIKASI KONFOUNDING). Pemilihan atau penentuan variabel independen yang masuk model seperti ini disebut dengan metode BACKWARD.
- 3). UJI ASUMSI REGRESI LINIER: Regresi linier merupakan salah satu perangkat statistik parametrik yang harus memenuhi berbagai macam asumsi, antara lain Normalitas, Homokedastisitas, Kolinearitas, Eksistensi, Linieritas, dan Independensi. Pemenuhan uji asumsi harus dilaporkan agar pembaca dapat pemahaman yang utuh tentang hasil analisis.

4). IDENTIFIKASI INTERAKSI: Setelah memperoleh model yang memuat variabel-variabel penting (signifikan atau konfounding), yang sudah memenuhi Uji ASUMSI, maka langkah terakhir adalah memeriksa kemungkinan adanya interaksi antar variabel independen. Penentuan variabel interaksi harus didasarkan pada pertimbangan logika substantif. Pengujian interaksi dilihat dari kemaknaan uji statistik. Bila variabel interaksi mempunyai nilai-p bermakna (signifikan $< 0,05$), maka variabel interaksi penting dimasukkan kedalam model.

5). FINAL MODEL: Model yang memuat variabel-variabel penting: signifikan, atau konfounding, atau interaksi

Ada beberapa metode yang dapat dilakukan untuk melakukan pemodelan dalam analisis multivariat regresi linier ganda, yaitu:

- a). ENTER, memasukkan semua variabel independen dengan serentak satu langkah, tanpa melewati kriteria kemaknaan statistik tertentu. Metode ini yang tepat dan sering digunakan, karena dalam pemodelan kita dapat melakukan pemilihan variabel independen berdasarkan pertimbangan aspek substansi, tidak hanya aspek statistik semata.
- b). FORWARD, memasukkan satu persatu variabel dari hasil pengkorelasiannya variabel dan memenuhi kriteria kemaknaan statistik untuk masuk ke dalam model, sampai semua variabel yang memenuhi kriteria tersebut masuk ke dalam model. Variabel yang masuk pertama kali adalah variabel yang mempunyai korelasi parsial terbesar dengan variabel dependen. Korelasi parsial adalah korelasi antara variabel independen dengan dependen, kriteria variabel yang dapat masuk bisa di setting, misalnya 0,10 atau 0,05 artinya variabel yang dapat masuk model bila variabel tersebut mempunyai nilai-P lebih kecil atau sama dengan 0,10 atau 0,05.
- c). BACKWARD, memasukkan semua variabel ke dalam model, tetapi kemudian satu persatu variabel independen dikeluarkan dari model berdasarkan kriteria kemaknaan tertentu, variabel yang pertama kali dikeluarkan adalah variabel yang mempunyai korelasi parsial terkecil dengan variabel dependen. Kriteria pengeluaran adalah 0,10 atau 0,05, artinya variabel yang mempunyai nilai-P lebih besar atau sama dengan 0,10 atau 0,05 dikeluarkan dari model.
- d). STEPWISE, model ini merupakan kombinasi antara metode backward dan Forward. Seperti halnya forward, metode Stepwise dimulai dari tanpa variabel sama sekali di dalam model. Lalu

satu variabel hasil pengkorelasian variabel dimasukkan ke dalam model. Lalu satu persatu variabel hasil pengkorelasian dimasukkan ke dalam model dan dikeluarkan dari model dengan kriteria tertentu. Variabel yang pertama masuk sama dengan metode forward yakni variabel yang mempunyai korelasi parsial terbesar. Selanjutnya setelah masuk, variabel pertama ini diperiksa lagi apakah harus dikeluarkan dari model menurut kriteria pengeluaran seperti metode backward.

5. Uji Diagnostik Asumsi Regresi Linier Multivariat

- a). Melakukan pengujian terhadap lima asumsi. Normalitas, Homokedastisitas, Kolinearitas, Eksistensi, dan Independensi.
- b). Melakukan pengujian adanya kolinearitas. Kolinearitas terjadi bila antar variabel independen terjadi saling hubungan yang kuat. Untuk mengetahui adanya kolinearitas dapat dilihat dari nilai koefisien korelasi-R, bila nilai r lebih tinggi dari 0,8 maka terjadi kolinearitas. Selain itu dapat diketahui dari nilai VIF atau tolerance, bila nilai VIF > 10, atau tolerance sekitar 1 (satu) maka model terjadi kolinearitas.
- c). Melakukan analisis interaksi. Setelah memperoleh model yang memuat variabel-variabel penting, maka langkah selanjutnya adalah memeriksa adanya interaksi antar variabel independen. Interaksi merupakan keadaan dimana hubungan antara satu variabel independen dengan dependen berbeda menurut tingkat variabel independen yang lain.
- d). Penilaian reliabilitas model. Model regresi yang sudah terpilih perlu dicek reliabilitasnya dengan cara membagi (split) sampel ke dalam dua kelompok. Untuk masing-masing sampel dibuat model dengan variabel yang sama, kemudian bandingkan antara model 1 dan model 2, bila hasilnya sama/hampir sama maka model regresi reliabel. Bila model reliabel maka seluruh sampel dapat digunakan untuk pembuatan model.

6. Aplikasi Pemodelan Regresi Linier Multivariat

Sebagai latihan kita melakukan analisis penelitian “faktor-faktor yang berhubungan dengan berat badan bayi”. Gunakan/aktifkan file data BBLR.189.SAV. Variabel independennya meliputi berat badan ibu dlm pounds (BWT), umur ibu (AGE), riwayat hipetensi (HT), riwayat merokok (SMOKE), frekuensi mengalami prematur (PTL) dan frekuensi melakukan ANC (FTV). Variabel dependennya berat badan bayi (BWT).

Kode variabel pada file data: LBW.SAV

Nama	Definisi Operasional	Hasil Ukur
Id	Nomor Identitas	
Low	Kondisi bayi dalam klasifikasi BBLR	0 = \geq 2500 g 1 = < 2500 g
Age	Umur ibu	tahun
Lwt	Berat ibu pada saat menstruasi terakhir	pounds
Race	Suku bangsa/ras	1= putih 2= hitam 3 = lainnya
Smoke	Kebiasaan merokok selama hamil	0 = tidak 1 = ya
Ptl	Riwayat mengalami prematur	0,1, 2, dst
Ht	Riwayat menderita hipertensi	0 = tidak 1 = ya
Ui	Terjadi/mengalami iritability Uterine	0 = tidak 1 = ya
Ftv	Frekuensi periksa hamil pada trimester pertama	0 ,1, 2 dst..
Bwt	Berat badan bayi	Gram

Data selengkapnya ada di lampiran:

6.1. Seleksi Bivariat

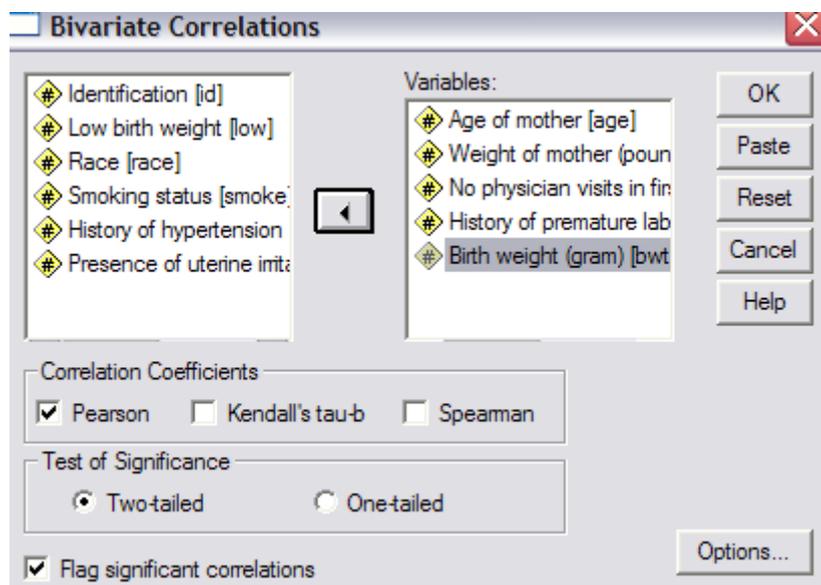
Seleksi bivariat masing-masing variabel independen dengan variabel dependen. Variabel yang dapat masuk model multivariat adalah variabel yang pada analisis bivariatnya mempunyai nilai p (p value) $< 0,25$. Namun ketentuan p value $< 0,25$ ini tidaklah harus dipenuhi manakala dijumpai ada suatu variabel yang walaupun p value-nya $> 0,25$ karena secara substansi sangat penting berhubungan dengan variabel dependen, maka variabel tersebut dapat diikuti dalam model multivariat.

Uji yang digunakan pada analisis bivariat tergantung dari variabel yang digunakan, bila variabel independennya numerik \rightarrow uji korelasi, bila independennya katagorik \rightarrow uji t atau uji anova.

a. **Bivariat uji korelasi:** melakukan analisis bivariat untuk variabel independen berjenis numerik: variabel berat badan ibu, umur ibu, frekuensi prematur, frekuensi anc:

Langkahnya:

1. Klik 'Analysis', sorot ke 'Correlate', sorot dan klik 'Bivariate'
2. Muncul dilayar menu 'Bivariate Correlations'
3. Pada kotak Variables, isikan semua variabel numerik baik untuk variabel independen (age,lwt,ptl,ftv) dan dependen (bwt)



4. Klik tombol 'OK'

Muncul dilayar output hasil sebagai berikut:

Correlations

Correlations

		Age of mother	Weight of mother (pounds)	No physician visits in first trimester	History of premature labor	Birth weight (gram)
Age of mother	Pearson Correlation	1	.180*	.215**	.072	.090
	Sig. (2-tailed)		.013	.003	.328	.219
	N	189	189	189	189	189
Weight of mother (pounds)	Pearson Correlation	.180*	1	.141	-.140	.186*
	Sig. (2-tailed)	.013		.054	.055	.010
	N	189	189	189	189	189
No physician visits in first trimester	Pearson Correlation	.215**	.141	1	-.044	.058
	Sig. (2-tailed)	.003	.054		.544	.426
	N	189	189	189	189	189
History of premature labor	Pearson Correlation	.072	-.140	-.044	1	-.155*
	Sig. (2-tailed)	.328	.055	.544		.034
	N	189	189	189	189	189
Birth weight (gram)	Pearson Correlation	.090	.186*	.058	-.155*	1
	Sig. (2-tailed)	.219	.010	.426	.034	
	N	189	189	189	189	189

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

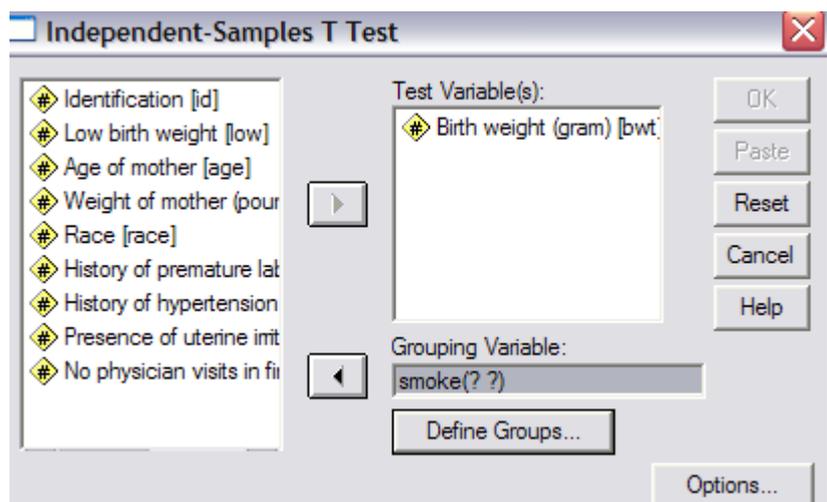
Hasil dari analisis bivariat dengan korelasi didapatkan nilai p value untuk variabel umur ($p=0,219$), berat badan ($p=0,010$), frekuensi anc ($p=0,426$), frekuensi prematur ($p=0,034$). Dari hasil ini dapat kita simpulkan bahwa variabel umur, berat badan dan frekuensi prematur mempunyai p value $< 0,25$, dengan demikian ketiga variabel tersebut dapat lanjut masuk ke pemodelan multivariat. Sedangkan untuk variabel frekuensi anc mempunyai p value $> 0,25$ (yaitu $p=0,426$) sehingga tidak bisa masuk ke multivariat, namun demikian oleh karena secara substansi frekuensi anc merupakan faktor yang sangat penting mempengaruhi berat badan bayi, maka variabel frekuensi anc tetap diikutkan dalam analisis multivariat.

b. Bivariat uji t: melakukan analisis bivariat untuk variabel independen berjenis katagorik: merokok dan riwayat hipertensi

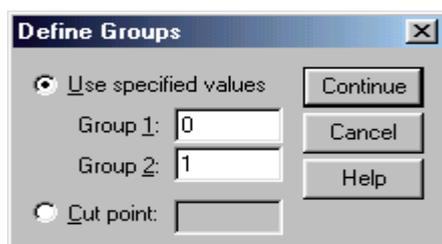
1. Merokok

Langkahnya:

1. Dari menu utama SPSS, pilih menu 'Analyze', kemudian pilih sub menu "Compare Means", lalu pilih "Independent-Samples T Test"
2. Pada layar tampak kotak yang di dalamnya ada kotak 'Test variable' dan 'Grouping Variable'. Ket: kotak *test variables tempat memasukkan variabel numeriknya*, sedangkan kotak *grouping variable* untuk memasukkan variabel katagoriknya, ingat jangan sampai terbalik.
3. Klik 'bwt' dan msukkan ke kotak 'Test variable'
4. Klik variabel 'smoke' dan masukkan ke kotak 'Grouping Variable'.



5. Klik 'Define Group', kemudian di layar nampak kotak isian. Anda diminta mengisi kode variabel 'smoke' ke dalam kedua kotak. Pada contoh ini, kita tahu bahwa '0' tidak merokok dan kode '1' untuk Yang merokok. Jadi ketiklah 0 pada *Group 1*" dan 1 pada "*Group 2*"



1. Klik "Continue"
2. Klik "OK" untuk menjalankan prosedur perintahnya, dan hasilnya sebagai berikut:

T-Test

Group Statistics

		Smoking status	N	Mean	Std. Deviation	Std. Error Mean
Birth weight (gram)	No		115	3054.96	752.409	70.163
	Yes		74	2773.24	660.075	76.732

Independent Samples Test

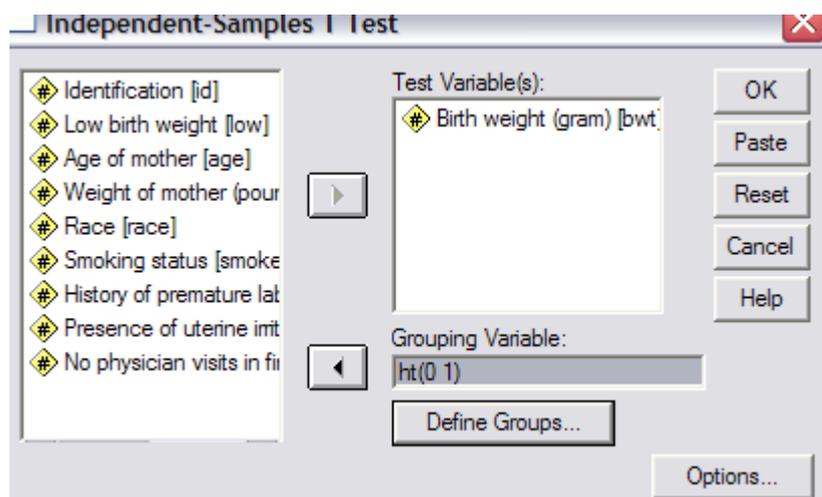
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Birth weight (gram)	Equal variances assumed	1.508	.221	2.634	187	.009	281.713	106.969	70.693	492.7
	Equal variances not assumed			2.709	170.0	.007	281.713	103.974	76.467	487.0

Hasil analisis hubungan merokok dengan berat bayi menghasilkan p value = 0,009, dengan demikian p value yang dihasilkan < 0,25 maka variabel merokok dapat lanjut ke multivariat.

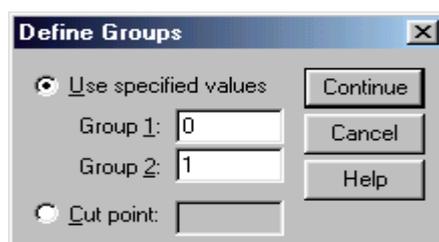
2. Riwayat Hipertensi

Langkahnya:

1. Dari menu utama SPSS, pilih menu 'Analyze', kemudian pilih sub menu "Compare Means", lalu pilih "Independent-Samples T Test"
2. Pada layar tampak kotak yang di dalamnya ada kotak 'Test variable' dan 'Grouping Variable'. Ket: kotak test variables tempat memasukkan variabel numeriknya, sedangkan kotak grouping variable untuk memasukkan variabel katagoriknya, ingat jangan sampai terbalik.
3. Klik 'bwt' dan msukkan ke kotak 'Test variable'
4. Klik variabel 'ht' dan masukkan ke kotak 'Grouping Variable'. (variabel yang sebelumnya (variabel smoke) dikeluarkan dahulu baru 'ht' dimasukkan



5. Klik 'Define Group', kemudian di layar nampak kotak isian. Anda diminta mengisi kode variabel 'smoke' ke dalam kedua kotak. Pada contoh ini, kita tahu bahwa '0' tidak ada hipertensi dan kode '1' ada hipertensi'. Jadi ketiklah 0 pada Group 1" dan 1 pada "Group 2"



6. Klik "Continue"

7. Klik "OK" untuk menjalankan prosedur perintahnya, dan hasilnya sebagai berikut:

Group Statistics

History of hypertension		N	Mean	Std. Deviation	Std. Error Mean
Birth weight (gram)	No	177	2972.31	709.226	53.309
	Yes	12	2536.75	917.341	264.813

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Birth weight (gram)	Equal variances assumed	1.419	.235	2.019	187	.045	435.56	215.709	10.024	861.1
	Equal variances not assumed			1.612	11.908	.133	435.56	270.126	-153.5	1025

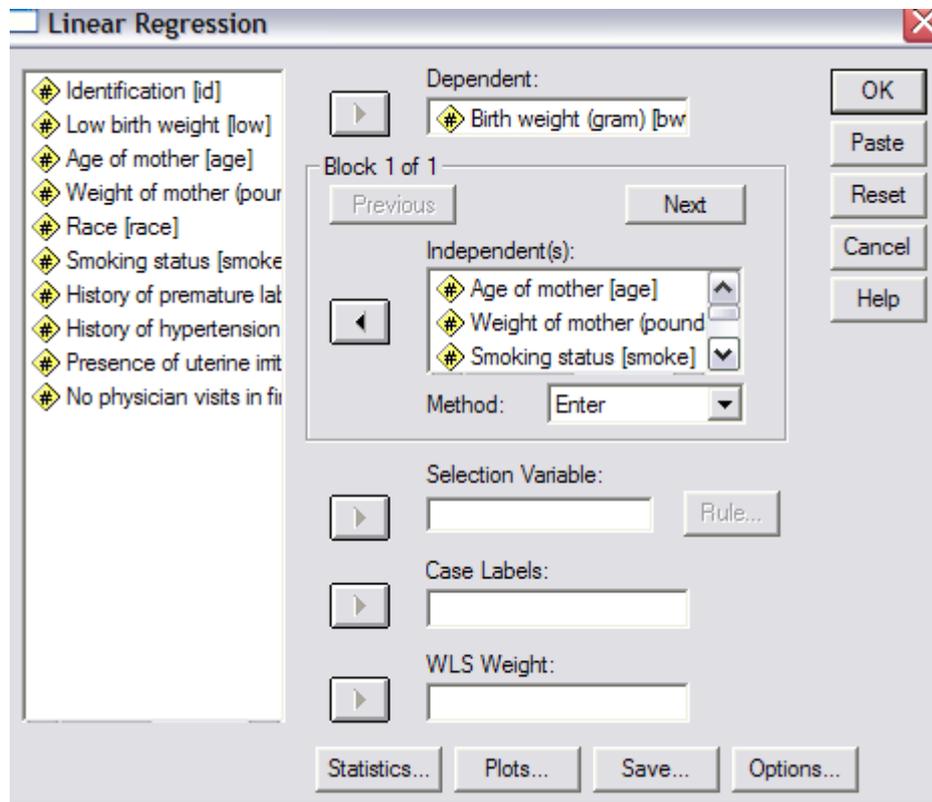
Dari hasil analisis bivariat uji t antara variabel riwayat adanya hipertensi dengan berat bayi didapatkan p value = 0,045, berarti p valuenya < 0,25 sehingga variabel riwayat adanya hipertensi dapat lanjut ke analisis multivariat

Dengan demikian selesailah sudah seleksi semua variabel independen, dari 6 variabel independen semuanya masuk ke proses berikutnya yaitu ke analisis multivariat.

6.2. Pemodelan Regresi Linier Multivariat

Setelah tahap bivariat selesai, tahap berikutnya melakukan analisis multivariat secara bersama-sama. Variabel yang valid dalam model multivariat adalah variabel yang mempunyai p value < 0,05. Bila dalam model multivariat dijumpai variabel yang Nilai-p nya > 0,05, maka variabel tersebut harus dikeluarkan dalam model. Pengeluaran variabel dilakukan tidak serempak, melainkan bertahap satu per satu dikeluarkan dimulai dari p value yang terbesar. Adapun proses selengkapnya sebagai berikut:

1. Klik 'Analysis', sorot 'Regression', sorot dan klik 'Linier' lalu muncul menu regresi linier,
 - a. Pada kotak 'dependen' isikan variabel dependen (dalam hal ini berarti bwt) dan kotak 'independen' isikan variabel independennya (dalam hal ini age, lwt, smoke, ht, pti, ftv)



5. Pada kotak 'Method', pilih Enter'
6. Abaikan lainnya
7. Klik 'OK', dan hasilnya

Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.340 ^a	.116	.086	696.829

- a. Predictors: (Constant), No physician visits in first trimester, Smoking status, History of hypertension, History of premature labor, Age of mother, Weight of mother (pounds)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11543236	6	1923872.611	3.962	.001 ^a
	Residual	88373817	182	485570.423		
	Total	99917053	188			

a. Predictors: (Constant), No physician visits in first trimester, Smoking status, History of hypertension, History of premature labor, Age of mother, Weight of mother (pounds)

b. Dependent Variable: Birth weight (gram)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2315.862	299.442		7.734	.000
	Age of mother	7.162	10.022	.052	.715	.476
	Weight of mother (pounds)	4.793	1.777	.201	2.698	.008
	Smoking status	-232.253	105.928	-.156	-2.193	.030
	History of premature labor	-154.002	106.574	-.104	-1.445	.150
	History of hypertension	-574.230	215.481	-.193	-2.665	.008
	No physician visits in first trimester	-2.847	49.705	-.004	-.057	.954

a. Dependent Variable: Birth weight (gram)

Dari kotak ‘Model Summary’ didapatkan nilai R Square sebesar 0,116, artinya keenam variabel independen dapat menjelaskan variabel berat bayi sebesar 11,6 % sedangkan sisanya dijelaskan oleh variabel lain. Dari hasil uji statistik (lihat kotak anova) didapatkan p value = 0,001 berarti persamaan garis regresi secara keseluruhan sudah signifikan. Namun demikian prinsip pemodelan harus yang sederhana variabelnya sehingga masing-masing variabel independen perlu di cek nilai p valuenya, variabel yang p valuenya > 0,05 dikeluarkan dari model. Ternyata dari 6 variabel independen (lihat kolom sig di kotak Coefficients) ada 3 variabel yang p valuenya > 0,05, yaitu umur (age) p=0,476, riwayat prematur (history prematur) p=0,150 dan frekuensi anc (no physician) p=0,954. Tahap berikutnya mengeluarkan variabel yang p valuenya > 0,05, pengeluaran variabel dimulai dari p value yang terbesar. Dengan demikian variabel yang kita coba keluarkan adalah frekuensi anc (No physician..).

Langkahnya:

1. Klik 'Analysis', sorot 'Regression', sorot dan klik 'Linier'
2. Di layar nampak pada kotak Dependen masih terisi 'bwt' lewati dan biarkan saja. Pada kotak Independen juga masih lengkap ada 6 variabel, namun sekarang anda harus keluarkan variabel 'no physician' dan masukkan ke kotak Variable di sebelah kiri.
3. Klik OK, dan hasilnya sebagai berikut:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.340 ^a	.116	.091	694.929

a. Predictors: (Constant), History of hypertension, Smoking status, Age of mother, History of premature labor, Weight of mother (pounds)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2317.608	297.074		7.801	.000
	Age of mother	7.051	9.807	.051	.719	.473
	Weight of mother (pounds)	4.781	1.759	.201	2.718	.007
	Smoking status	-232.224	105.638	-.156	-2.198	.029
	History of premature labor	-153.747	106.191	-.104	-1.448	.149
	History of hypertension	-573.011	213.841	-.192	-2.680	.008

a. Dependent Variable: Birth weight (gram)

6.3. Uji Konfounding

Setelah variabel frekuensi anc dikeluarkan, kita cek dulu apakah setelah dikeluarkan, ada perubahan besar(berubah lebih dari 10 %) untuk R Square dan Coef. B. Bila ada perubahan yang besar maka variabel tersebut tidak jadi dikeluarkan dalam model (tetap dipertahankan di model). Untuk nilai R Square ternyata tidak ada perubahan yaitu tetap 0,116. Sedangkan untuk **coefisien B**, Sekarang kita bandingkan nilai coefisien B untuk variabel umur, berat ibu, merokok, riwayat prematur dan riwayat hipertensi antara sebelum dan sesudah variabel frekuensi anc dikeluarkan, hasil perhitungannya sebagai berikut:

Variabel	Anc msih ada	Anc dikeluarkan	perubahan Coef.
Age	7,1	7,0	1,4 %
Lwt	4,7	4,7	0 %
smoke	-232,2	-232,2	0 %
ptl	-154,0	153,7	0,1 %
hi	-574,2	573,0	0,1 %
ftv	-2,8	-	

Dari perhitungan perubahan nilai coefisien B pada masing-masing variabel, ternyata tidak ada yang berubah lebih dari 10 %, dengan demikian variabel frekuensi anc kita keluarkan dari model.

Selanjutnya kita lihat kembali bahwa pada model masih ada variabel yang p value > 0,05. Sekarang kita akan keluarkan variabel umur (p value =0,473).

Langkah/proses:

1. Klik 'Analysis', sorot 'Regression', sorot dan klik 'Linier'
2. Di layar nampak pada kotak Dependen masih terisi 'bwt' lewati dan biarkan saja. Pada kotak Independen juga masih terisi ada 5 variabel, namun sekarang anda harus keluarkan variabel 'umur (age)' dan masukkan ke kotak Variable di sebelah kiri.
3. Klik OK, dan hasilnya sebagai berikut:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.336 ^a	.113	.094	694.016

- a. Predictors: (Constant), History of hypertension, Smoking status, History of premature labor, Weight of mother (pounds)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2449.121	233.779		10.476	.000
	Weight of mother (pounds)	5.035	1.721	.211	2.925	.004
	Smoking status	-236.420	105.338	-.159	-2.244	.026
	History of premature labor	-145.412	105.417	-.098	-1.379	.169
	History of hypertension	-582.566	213.148	-.195	-2.733	.007

a. Dependent Variable: Birth weight (gram)

Setelah variabel umur dikeluarkan, nilai R Square ternyata ada sedikit perubahan yaitu menjadi 0,113. sedangkan untuk **coefisian B**, hasil perhitungannya sebagai berikut:

Variabel	Masih lengkap	umur dikeluarkan	perubahan Coef.
Age	7,1	-	-
Lwt	4,7	5,0	6,3 %
smoke	-232,2	-236,4	1,8 %
ptl	-154,0	145,4	6,1 %
hi	-574,2	582,5	1,3 %
ftv	-2,847	-	

Dari hasil perhitungan perubahan coef. Ternyata tidak ada yang lebih dari 10 %, dengan demikian variabel umur kita keluarkan dari model.

Langkah selanjutnya mengeluarkan variabel Riwayat mengalami prematur, Prosesnya/langkahnya sama dengan diatas, Klik Analysis, sorot Regression, ..dst. Pada **kotak independen** variabel riwayat mengalami prematur dikeluarkan dan dimasukkan ke kotak variable disebelah kiri, dan hasilnya sebagai berikut:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.322 ^a	.104	.089	695.707

a. Predictors: (Constant), History of hypertension, Smoking status, Weight of mother (pounds)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2390.105	230.391		10.374	.000
	Weight of mother (pounds)	5.352	1.710	.224	3.130	.002
	Smoking status	-263.009	103.812	-.177	-2.534	.012
	History of hypertension	-586.722	213.646	-.197	-2.746	.007

a. Dependent Variable: Birth weight (gram)

Hasil R Square turun sedikit yaitu menjadi 0,104. Sedangkan hasil perhitungan perubahan Coef. B dapat dilihat sebagai berikut:

Variabel	Masih lengkap	Prematur keluar	perubahan Coef.
Age	7,1	-	-
Lwt	4,7	5,3	12,3 %
smoke	-232,2	-236,4	1,7 %
ptl	-154,0	-	-
hi	-574,2	582,5	1,3 %
ftv	-2,847	-	-

Hasil perhitungan setelah dikeluarkan variabel prematur, ternyata coefisin B pada variabel berat badan ibu (lwt) beubah sebesar 12,3 % dengan demikian variabel riwayat mengalami prematur tidak jadi dikeluarkan dan tetap dipertahankan dalam model multivariat. Dari hasil analisis ternyata tidak

ada lagi yang p value-nya > 0,05 dengan demikian proses pencarian variabel yang masuk dalam model telah selesai dan model yang terakhir adalah sebagai berikut:

Model Summary^a

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.336 ^a	.113	.094	694.016	.222

- a. Predictors: (Constant), History of premature labor, History of hypertension, Smoking status, Weight of mother (pounds)
 b. Dependent Variable: Birth weight (gram)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2449.121	233.779		10.476	.000		
	Weight of mother (pounds)	5.035	1.721	.211	2.925	.004	.925	1.081
	Smoking status	-236.420	105.338	-.159	-2.244	.026	.964	1.037
	History of hypertension	-582.566	213.148	-.195	-2.733	.007	.943	1.060
	History of premature labor	-145.412	105.417	-.098	-1.379	.169	.947	1.056

- a. Dependent Variable: Birth weight (gram)

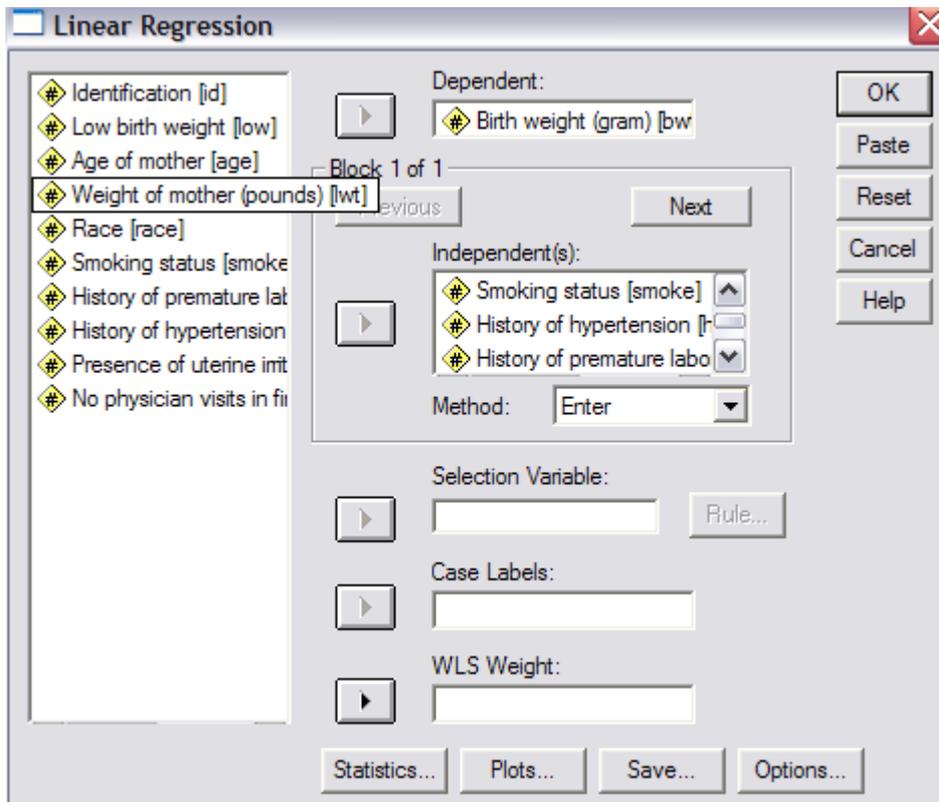
6.4. Uji Asumsi Regresi Linier Multivariat

Agar persamaan garis yang digunakan untuk memprediksi menghasilkan angka yang valid, maka persamaan yang dihasilkan harus memenuhi asumsi-asumsi yang dipersyaratkan uji regresi linier ganda. Adapun uji asumsinya sebagai berikut:

Langkahnya:

1. Klik 'Analysis', sorot 'Regression', sorot dan klik 'Linier'
2. Masukkan dalam kotak Dependen variabel 'bwt'

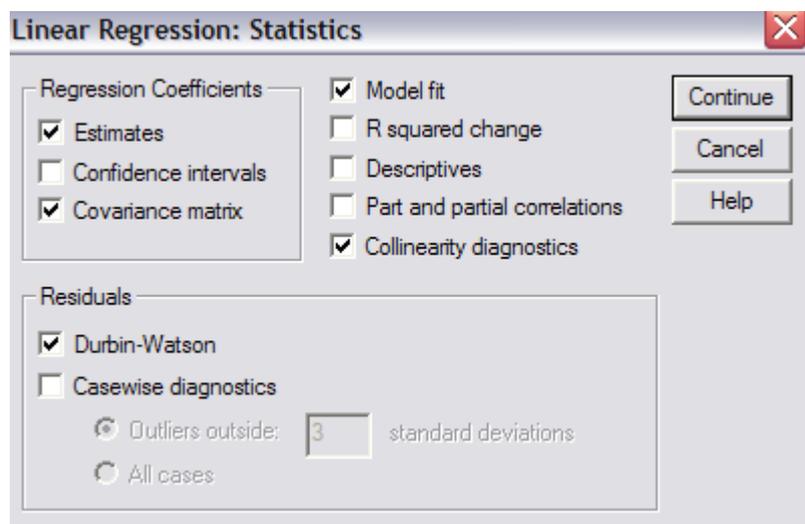
3. Masukkan dalam kotak Independen variabel berat badan ibu (lwt), merokok(smoke), riwayat hipertensi (hi) dan variabel riwayat prematur(ptl)



4. Klik tombol Statistics

5. Klik kotak 'Collinearity diagnostic' dan klik kotak 'Covariance matrix' (perintah ini untuk uji asumsi multicolinearity)

6. Klik kotak 'Durbin-Watson' (perintah ini untuk uji asumsi Independensi)

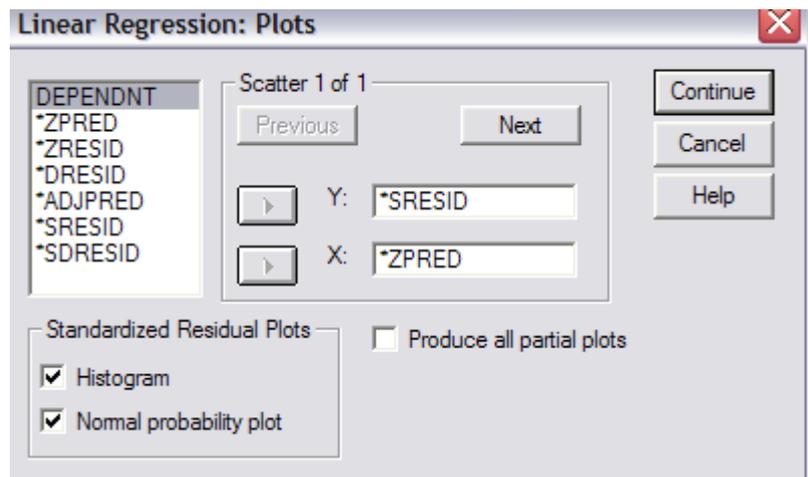


7. Klik Continue

8. Klik tombol 'Plot'

9. Masukkan 'SRESID' ke kotak Y, dan masukan 'ZPRED' ke kotak X (perintah ini untuk uji asumsi Homoscedasity)

10. Klik kotak 'histogram' dan kotak 'Normal probability plot' (perintah ini untuk uji asumsi Normality)



11. Klik Continue

Hasilnya:

a. **Asumsi Eksistensi** (Variabel Random)

Untuk tiap nilai dari variabel X (variabel independen), variabel Y (dependen) adalah variabel random yang mempunyai mean dan varian tertentu. Asumsi ini berkaitan dengan teknik pengambilan sampel. Untuk memenuhi asumsi ini, sampel yang diambil harus dilakukan secara random. Cara mengetahui asumsi eksistensi dengan cara melakukan analisis deskriptif variabel residual dari model, bila residual menunjukkan adanya mean mendekati nilai nol dan ada sebaran (varian atau standar deviasi) maka asumsi eksistensi terpenuhi. Hasil analisis:

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2249.77	3602.03	2944.66	245.079	189
Std. Predicted Value	-2.835	2.682	.000	1.000	189
Standard Error of Predicted Value	67.193	292.804	103.399	45.407	189
Adjusted Predicted Value	1955.43	3616.97	2943.73	251.196	189
Residual	-2082.610	1921.631	.000	686.593	189
Std. Residual	-3.001	2.769	.000	.989	189
Stud. Residual	-3.015	2.782	.001	1.005	189
Deleted Residual	-2102.316	1940.423	.923	708.619	189
Stud. Deleted Residual	-3.084	2.835	.000	1.010	189
Mahal. Distance	.768	32.469	3.979	5.320	189
Cook's Distance	.000	.209	.007	.019	189
Centered Leverage Value	.004	.173	.021	.028	189

a. Dependent Variable: Birth weight (gram)

Hasil dari output diatas menunjukkan angka residual dengan mean 0,000 dan standar deviasi 686,59. Dengan demikian asumsi Eksistensi terpenuhi

b. Asumsi Independensi

Suatu keadaan dimana masing-masing nilai Y bebas satu sama lain. Jadi nilai dari tiap-tiap individu saling berdiri sendiri. Tidak diperbolehkan nilai observasi yang berbeda yang diukur dari satu individu diukur dua kali. Untuk mengetahui asumsi ini dilakukan dengan cara mengeluarkan uji Durbin Watson, bila nilai Durbin -2 s.d. $+2$ berarti asumsi independensi terpenuhi, sebaliknya bila nilai Durbin < -2 atau $> +2$ berarti asumsi tidak terpenuhi

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.336 ^a	.113	.094	694.016	.222

a. Predictors: (Constant), History of premature labor, History of hypertension, Smoking status, Weight of mother (pounds)

b. Dependent Variable: Birth weight (gram)

Dari hasil uji didapatkan koefisien Durbin Watson 0,222, berarti asumsi independensi terpenuhi.

c. Asumsi Linieritas

Nilai mean dari variabel Y untuk suatu kombinasi $X_1, X_2, X_3, \dots, X_k$ terletak pada garis/bidang linier yang dibentuk dari persamaan regresi. Untuk mengetahui asumsi linieritas dapat diketahui dari uji ANOVA (overall F test) bila hasilnya signifikan ($p \text{ value} < \alpha$) maka model berbentuk linier. Hasil uji asumsi:

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11291987	4	2822996.778	5.861	.000 ^a
	Residual	88625066	184	481657.965		
	Total	99917053	188			

a. Predictors : (Constant), History of premature labor, History of hypertension, Smoking status, Weight of mother (pounds)

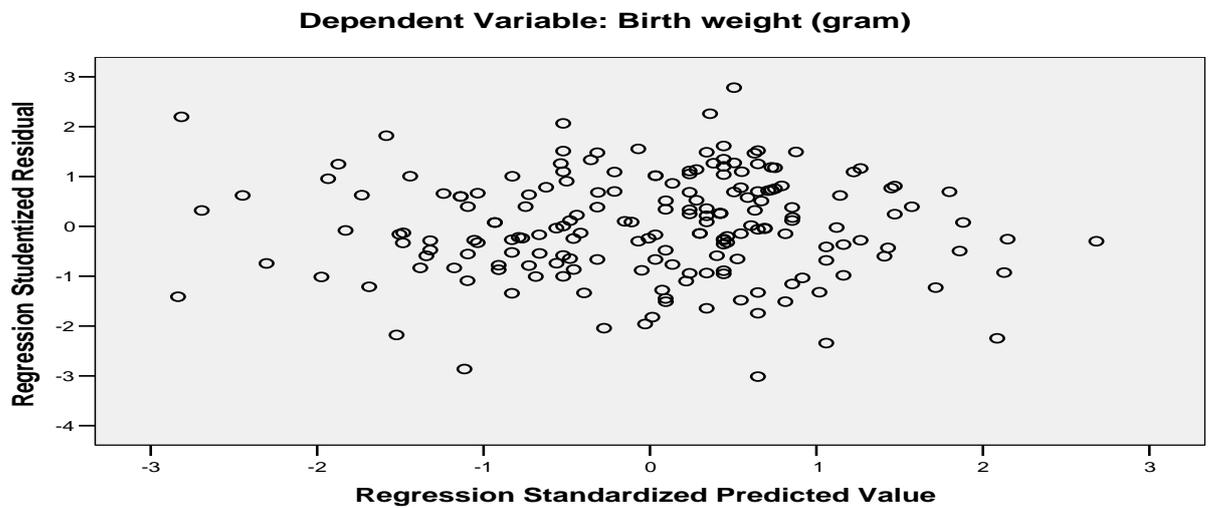
b. Dependent Variable: Birth weight (gram)

Dari output diatas menghasilkan uji anova 0,0005, berarti asumsi linearitas terpenuhi

d. Asumsi Homoscedasticity

Varian nilai variabel Y sama untuk semua nilai variabel X. Homoscedasticity dapat diketahui dengan melakukan pembuatan plot residual. Bila titik tebaran tidak berpola tertentu dan menyebar merata disekitar garis titik nol maka dapat disebut varian homogen pada setiap nilai X dengan demikian asumsi homoscedasticity terpenuhi. Sebaliknya bila titik tebaran membentuk pola tertentu misalnya mengelompok di bawah atau di atas garis tengah nol, maka diduga variannya terjadi heteroscedasticity.

Scatterplot

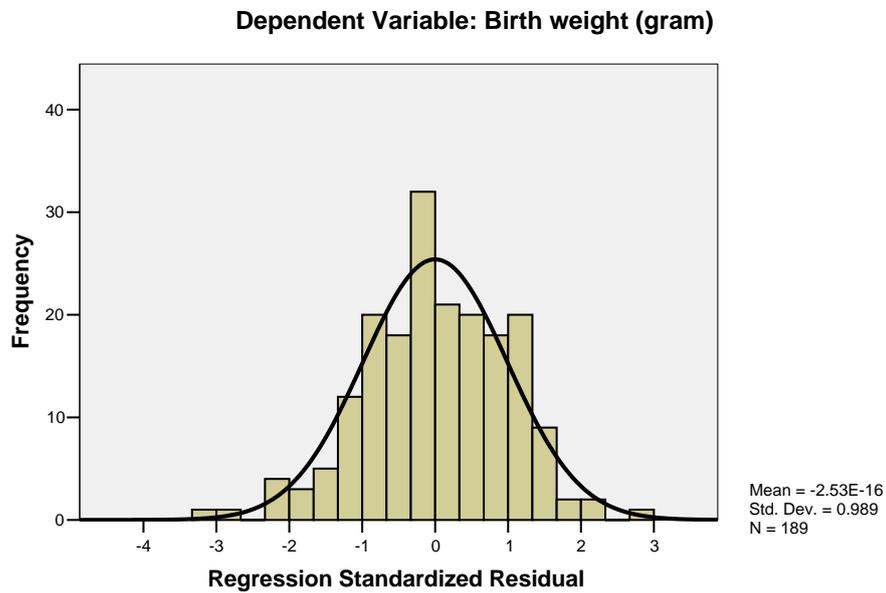


Dari hasil plot diatas terlihat tebaran titik mempunyai pola yang sama antara titik-titik diatas dan dibawah garis diagonal 0. Dengan demikian asumsi homoscedasity terpenuhi

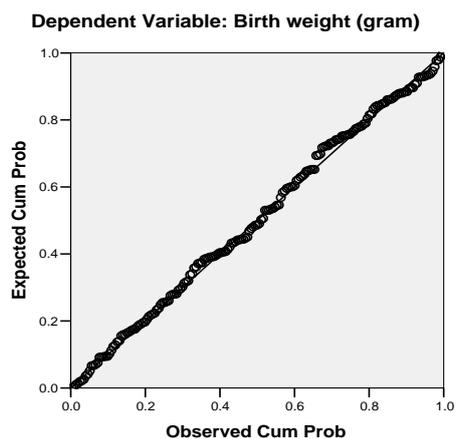
e. Asumsi Normalitas

Variabel Y mempunyai distribusi normal untuk setiap pengamatan variabel X. dapat diketahui dari Normal P-P Plot residual, bila data menyebar di sekitar garis diagonal dan mengikuti arah garis diagonal, maka model regresi memenuhi asumsi model regresi tidak memenuhi asumsi normalitas.

Histogram



Normal P-P Plot of Regression Standardized Residual



Dari grafik histogram dan grafik normal P-P plot terbukti bahwa bentuk distribusinya normal, berarti asumsi normality terpenuhi.

f. Diagnostik Multicollinearity

Dalam regresi linier tidak boleh terjadi sesama variabel independen berkorelasi secara kuat (multicollinearity). Untuk mendeteksi collinearity dapat diketahui dari nilai VIF (variance inflation factor), bila nilai VIF lebih dari 10 maka mengindikasikan telah terjadi collinearity.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2449.121	233.779		10.476	.000		
	Weight of mother (pounds)	5.035	1.721	.211	2.925	.004	.925	1.081
	Smoking status	-236.420	105.338	-.159	-2.244	.026	.964	1.037
	History of hypertension	-582.566	213.148	-.195	-2.733	.007	.943	1.060
	History of premature labor	-145.412	105.417	-.098	-1.379	.169	.947	1.056

a. Dependent Variable: Birth weight (gram)

Dari hasil uji asumsi didapatkan nilai VIF tidak lebih dari 10, dengan demikian tidak ada Multicollinearity antara sesama variabel independen

Dari hasil uji asumsi dan uji kolinearitas ternyata semua asumsi terpenuhi sehingga model dapat digunakan untuk memprediksi berat badan bayi.

6.5. Uji Interaksi Regresi Linier

Langkah selanjutnya adalah membuat MODEL INTERAKSI, variabel yang secara logika substansi berinteraksi adalah ROKOK dengan HIPERTENSI. Berbeda dengan regresi logistik, dimana variabel interaksi bisa langsung masuk ke dalam pemodelan, sedangkan regresi linier harus dibuat variabel interaksi secara manual terlebih dahulu. Gunakan perintah compute untuk membuat variabel baru SMOKE.HT yang merupakan perkalian antara SMOKE dengan HT. Kemudian tambahkan variabel tersebut ke model regresi linier yang sudah diuji asumsinya, hasilnya sebagai berikut:

		Coefficients ^a				
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	2454.105	239.152		10.262	.000
	Berat ibu	5.004	1.750	.210	2.859	.005
	Ibu perokok	-239.669	110.046	-.161	-2.178	.031
	Hipertensi	-600.411	272.883	-.201	-2.200	.029
	Riwayat melahirkan prematur (kali)	-144.242	106.285	-.098	-1.357	.176
	Smoke.HT	45.314	430.848	.010	.105	.916

a. Dependent Variable: Berat bayi lahir (gr)

Ternyata variabel SMOKE dengan HT tidak berinteraksi, nilai-p = 0,916

6.6. Model Akhir dan Interpretasi

Langkah selanjutnya adalah membuat FINAL MODEL, model yang terakhir adalah model tanpa interaksi sebagai berikut:

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.336 ^a	.113	.094	694.016	.222

- a. Predictors: (Constant), History of premature labor, History of hypertension, Smoking status, Weight of mother (pounds)
 b. Dependent Variable: Birth weight (gram)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	2449.121	233.779		10.476	.000		
	Weight of mother (pounds)	5.035	1.721	.211	2.925	.004	.925	1.081
	Smoking status	-236.420	105.338	-.159	-2.244	.026	.964	1.037
	History of hypertension	-582.566	213.148	-.195	-2.733	.007	.943	1.060
	History of premature labor	-145.412	105.417	-.098	-1.379	.169	.947	1.056

- a. Dependent Variable: Birth weight (gram)

Interpretasi model:

Setelah dilakukan analisis, ternyata variabel independen yang masuk model regresi adalah berat badan ibu, ibu merokok, riwayat hipertensi, dan riwayat prematur. Pada tabel 'Model Summary' terlihat koefisien determinasi (R square) menunjukkan nilai 0,113 artinya bahwa model regresi yang diperoleh dapat menjelaskan 11,3 % variasi variabel dependen berat bayi. Atau dengan kata lain keempat variabel independen tsb dapat menjelaskan variasi variabel berat bayi sebesar 11,3 %.. Kemudian pada kotak 'ANOVA', kita lihat hasil uji F yang menunjukkan nilai P (sig) = 0,000, berarti pada alpha 5% kita dapat menyatakan bahwa model regresi cocok (fit) dengan data yang ada. Atau dapat diartikan kedua variabel tersebut secara signifikan dapat untuk memprediksi variabel berat bayi.

Pada kotak '*Coefficient*' kita dapat memperoleh persamaan garisnya, pada kolom B (di bagian *Variabel In Equation*) di atas, kita dapat mengetahui koefisien regresi masing-masing variabel. Dari hasil di atas, persamaan regresi yang diperoleh adalah

$$\text{Berat Bayi} = 2449,1 + 5,0 \text{ Lwt} - 236,4 \text{ smoke} - 582,5 \text{ Hi} - 145,4 \text{ Ptl}$$

Dengan model persamaan ini, kita dapat memperkirakan berat badan bayi dengan menggunakan variabel berat badan ibu, merokok dan hipertensi. Adapun arti koef. B untuk masing-masing variabel adalah sebagai berikut:

- Setiap kenaikan berat badan ibu sebesar 1 kg, maka berat badan bayi akan naik sebesar 5,0 gram setelah dikontrol variabel merokok, hipertensi dan prematur
- Pada ibu yang merokok berat bayinya akan lebih rendah sebesar 236,4 gram setelah dikontrol variabel berat badan, hipertensi dan prematur.
- Pada ibu yang menderita hipertensi, berat bayinya akan lebih rendah sebesar 582,5 gram setelah dikontrol variabel berat badan ibu, merokok dan prematur.

Kolom **Beta** dapat digunakan untuk mengetahui variabel mana yang paling besar perannya (pengaruhnya) dalam menentukan variabel dependennya (berat badan bayi). Semakin besar nilai beta semakin besar pengaruhnya terhadap variabel dependennya. Pada hasil di atas berarti variabel yang paling besar pengaruhnya terhadap penentuan berat badan bayi adalah berat badan ibu..

DAFTAR PUSTAKA

- Ariawan, Iwan 1998. *Besar dan metode Sampel pada Penelitian Kesehatan*, Jurusan Biostatistik dan Kependudukan Fakultas Kesehatan Masyarakat. Universitas Indonesia, Depok
- Chow SC, Shao J, Wang H. 2008. *Sample Size Calculations in Clinical Research*. Second edition. Chapman & Hall/CRC
- Daniel, Wayne W, 1999. *Biostatistics: A Foundation for Analysis in the Health Sciences*, John Wiley & Sons Inc.
- Kleimbaum, Kupper, Muller, 1988. *Applied Regression Analysis and Other Multivariable Method*. 2nd ed. Boston: PWS Kent Pub.Co.
- Kleinbaum, David G dan Klein, Mitchel, 2010. *Logistic Regression: A Self Learning Text*. Third edition. Springer-Verlag. New York Berlin Heidelberg
- Kusma, Jan W, 1984. *Basic Statistics for the Health Sciences*. California: Mayfield Publishing Company.
- Lemeshow S, Hosmer D, Klar J, Lwanga S.,1997. *Besar Sampel dalam Penelitian Kesehatan*. Gajah Mada University Press. Yogyakarta
- Murti, Bhisma, 1997. *Prinsip-prinsip Metode Riset Epidemiologi*. UGM Press, Yogyakarta
- Pagano, Marcello, Kimberlee Gauvreau, 1993. *Principles of Biostatistics*. Belmont: Duxbury Press.
- Sastroasmoro, S., 2002. *Dasar-dasar Metodologi Penelitian Klinis*. Binarupa Aksara, Jakarta.
- SPSS Inc, 2005. *SPSS Advanced Models 17*. SPSS Inc, USA
- Supranto, 2004. *Analisis Multivariat*, Rineka Cipta, Jakarta.
- Tabachnick, Barbara G.,2001. *Using Multivariate Statistics*. USA. A Pearson Education Company

Tentang Penulis

Lahir di Desa Tagog, Kecamatan Conggeang, Kab.Sumedang, Jawa Barat tanggal 28 Juli 1970. Menyelesaikan Diploma III Jurusan Gizi, di Politeknik Kesehatan Bandung pada Tahun 1992, Sarjana Kesehatan Masyarakat di Universitas Indonesia Jakarta tahun 2000, dan Magister Kesehatan Masyarakat di Universitas Indonesia Jakarta tahun 2006, serta menyelesaikan Doktor Kesehatan Masyarakat di Universitas Indonesia Jakarta tahun 2015.

Sejak lulus dari Politeknik Kesehatan Bandung pada tahun 1992, langsung bekerja sebagai dosen di Jurusan Gizi Politeknik Kesehatan Kementerian Kesehatan Bandung sampai sekarang. Sejak Tahun 2018 menjadi kepala Pusat Penelitian dan pengabdian kepada Masyarakat di Poltekkes Kemenkes Bandung.

Beliau mengajar di bidang Statistik Deskriptif dan Inferensial, termasuk Statistik Non-parametrik, Manajemen dan Analisis Data. Mengajar Metode Penelitian Kuantitatif, Rancangan Sampel, dan Aplikasi Analisis Multivariat pada berbagai jenis studi penelitian kesehatan dan memberikan bantuan teknis dan konsultasi di bidang Biostatistik, Metode Penelitian, dan Teknik Analisis Data di berbagai universitas dan institusi kesehatan di Indonesia pada masyarakat umum, akademisi dan mahasiswa DIII, D IV, S1, S2, dan S3.



Nama dan Gelar:
Dr. Rr. Nur Fauziah, SKM, MKM, RD

Alamat Kantor:
Politeknik Kesehatan Kemenkes Bandung
Jalan Pajajaran nomor 56 Bandung 40171
Telp : (022) 4231627, (022) 4231639, Fax : (022) 4231640
E-mail : info@poltekkesbandung.ac.id
Website : www.poltekkesbandung.ac.id

Jurusan Gizi
Phone: (+62-22) 6628150
Fax: (+62-22) 2000505
Hp: 0817226151
Email: roronur70@yahoo.com

Alamat Rumah:
Komplek Permata Biru, Blok W, No. 210, RT 09/20
Kel. Cinunuk, Kec. Cileunyi, Bandung 40393