



MODEL REGRESI BERGANDA DENGAN METODE BACKWARD ELIMINATION (STUDI KASUS : PENDAPATAN PAJAK DAERAH BEBERAPA KABUPATEN/KOTA DI JAWA TIMUR 2020)

Sigap Abror Falah^{1*}

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Malang,
sigap.abror.2003126@students.um.ac.id

*Email : sigap.abror.2003126@students.um.ac.id

Abstrak

Membayar pajak adalah kewajiban yang harus dilakukan setiap warga negara sesuai peraturan yang berlaku. Salah satu jenis pajak yang wajib ditunaikan oleh warga negara yaitu pajak daerah. Pajak daerah merupakan iuran wajib yang dipungut pemerintah daerah yang bersifat memaksa dan digunakan untuk menjalankan pemerintahan. Beberapa sumber menyatakan bahwa faktor yang mempengaruhi besarnya pajak daerah yaitu: Produk Domestik Regional Bruto (PDRB), Jumlah Penduduk, dan Inflasi. Dilakukan pengujian kebenaran pernyataan tersebut menggunakan model terbaik regresi linier/linear berganda dengan metode backward elimination yang dapat digunakan untuk memodelkan model regresi pada pajak daerah dengan akurat. Berdasarkan hasil pengujian, diperoleh model regresi terbaik menggunakan metode backward elimination dengan persamaan $Y = -194.7 + (0.008654)X_1$. Berdasarkan perhitungan menggunakan model yang telah dirumuskan, diperoleh kesimpulan bahwa faktor yang paling berpengaruh terhadap nilai pajak daerah di beberapa kabupaten/kota di Jawa Timur adalah Produk Domestik Regional Bruto (PDRB) dengan koefisien determinasi sebesar 0,9683 atau 96,83%.

Kata kunci: backward elimination, pajak daerah, PDRB, regresi linier berganda.

PENDAHULUAN

Berdasarkan UU Republik Indonesia No. 28 Tahun 2007, pajak adalah kontribusi wajib kepada negara yang terutang oleh orang pribadi atau badan yang bersifat memaksa berdasarkan Undang-Undang, dengan tidak mendapatkan imbalan secara langsung dan digunakan untuk keperluan negara bagi sebesar-besarnya kemakmuran rakyat. Pajak bersifat mengikat pada seluruh wilayah negara kesatuan Republik Indonesia yang terdiri dari 34 provinsi, dimana setiap provinsi terdiri dari beberapa kabupaten/kota. Dalam pengaturan perpajakan daerah di seluruh wilayah NKRI, setiap daerah berhak mengatur pajak daerah masing-masing. Hal ini juga berlaku untuk kabupaten/kota di Jawa Timur.

Pajak daerah merupakan iuran wajib yang dipungut oleh pemerintah daerah yang bersifat memaksa dan digunakan sebesar-besarnya untuk menjalankan pemerintahan. Berdasarkan penelitian (Rooy & Budiarmo, 2015), faktor-faktor yang berpengaruh signifikan terhadap pajak daerah yaitu Produk Domestik Bruto (PDRB), inflasi, dan jumlah penduduk. Namun demikian, belum dapat ditentukan faktor mana yang paling berpengaruh terhadap besaran pajak yang ditentukan oleh setiap daerah.

Salah satu metode statistika yang dapat digunakan untuk menyelesaikan masalah ini adalah dengan menggunakan analisis regresi linier. Analisis regresi merupakan analisis yang dapat digunakan untuk mengkaji pengaruh variabel independen terhadap variabel dependen dan meramalkan variabel dependen berdasarkan variabel independen (Ekanayake, Rankothge, Weliwatta, & Jayasinghe, 2021). Berdasarkan jumlah variabel independen yang digunakan, analisis regresi linier dapat dibagi menjadi dua jenis, yaitu analisis regresi linier sederhana dan analisis regresi linier berganda. Untuk menentukan faktor dominan dari besaran pajak di setiap daerah, dapat digunakan analisis regresi linier berganda yang melibatkan lebih dari satu variabel independen.

Terdapat dua metode dalam pembentukan model analisis regresi linier berganda, yaitu metode *backward elimination* dan metode *forward selection*. Pada penelitian ini, dilakukan perumusan model analisis regresi linier berganda dengan metode *backward elimination*. Metode *backward elimination* pernah digunakan untuk menentukan persamaan regresi linier berganda terhadap jumlah kecelakaan lalu lintas di kota Medan pada tahun 2014 (Novelysa, Siagian, & Bangun, 2014). Selain itu, metode *backward elimination* telah digunakan untuk menentukan persamaan regresi linier berganda terhadap dugaan tindak pidana di Kota Binjai pada tahun 2023 (Yulianti & Sembiring, 2023).

Berdasarkan latar belakang masalah tersebut, maka penelitian ini dilakukan untuk mengetahui model persamaan regresi linier berganda yang tepat dalam menentukan faktor yang paling berpengaruh terhadap

besaran pajak daerah di beberapa kabupaten/kota di Jawa Timur dengan menggunakan metode *backward elimination* dengan bantuan *software* Ms. Excel dan Minitab 19.

METODE

Data dan Variabel

Data yang digunakan dalam penelitian ini merupakan data sekunder yang di peroleh dari Badan Pusat Statistik (BPS) Jawa Timur tahun 2020. Sampel yang digunakan dalam penelitian ini berjumlah delapan data yang masing-masing merupakan data pajak daerah beberapa kabupaten/kota di Jawa Timur pada tahun 2020. Variabel dependen (Y) yang digunakan dalam penelitian ini adalah pajak daerah (djpk.kemenkeu.go.id, 2021). Variabel independen (X) yang digunakan yaitu: Produk Domestik Regional Bruto (X1), inflasi (X2), dan jumlah penduduk (X3). Data PDRB dan inflasi dikumpulkan melalui laman Badan Pusat Statistik Provinsi Jawa Timur (jatim.bps.go.id, 2021). Data jumlah penduduk diperoleh dari laman resmi Badan Pusat Statistik (bps.go.id, 2021).

Prosedur

Penelitian ini dilaksanakan menggunakan bantuan *software* Ms. Excel dan Minitab 19. Prosedur penelitian yang dilakukan dapat diuraikan sebagai berikut:

Pengujian Multikolinearitas

Suatu model regresi terbaik merupakan model regresi yang tidak memiliki kasus multikolinearitas. Pengujian ini menggunakan nilai Variance Inflation Factor (VIF), yaitu jika nilai VIF kurang dari 10, maka tidak terdapat kasus multikolinearitas pada model regresi. Sebaliknya, jika VIF lebih dari 10, maka terdapat kasus multikolinearitas antar variabel bebas pada model tersebut. Namun demikian, walaupun tidak terdapat nilai VIF yang lebih dari 10, tidak menjamin bahwa tidak ada kasus multikolinearitas. Sehingga, perlu diperhatikan kembali apakah terdapat perubahan tanda dari koefisien setiap variabel bebas dengan persentase korelasi setiap variabel bebas terhadap variabel terikat. Ketika terdapat perubahan tanda, maka dapat disimpulkan bahwa terjadi kasus multikolinearitas, serta berlaku sebaliknya.

Penentuan Model Regresi Linear dengan Metode Backward Elimination

Metode *backward elimination* diawali dengan memasukkan semua variabel bebas pada model, kemudian dieliminasi satu persatu dengan menggunakan Uji F dan Uji t. Uji F dilakukan untuk mengetahui apakah semua variabel bebas berpengaruh secara signifikan terhadap variabel terikat (Sulistiyono & Sulistiyowati, 2017). Nilai F_{hitung} diperoleh menggunakan *analysis of variance* menggunakan *software* minitab 19. Uji t dilakukan untuk mengetahui apakah dalam model regresi, variabel bebas secara parsial berpengaruh signifikan terhadap variabel terikat (Sulistiyono & Sulistiyowati, 2017).

Pengujian Asumsi Klasik Regresi Linier

Uji normalitas

Uji normalitas dilakukan untuk mengetahui apakah data berdistribusi normal atau tidak. Uji normalitas dapat dilakukan menggunakan Kolmogorof-Smirnov test. Jika p-value bernilai lebih dari 0,05, maka data berdistribusi normal.

Uji autokorelasi

Uji autokorelasi dilakukan untuk mengetahui adanya korelasi yang sempurna antar variabel. Deteksi uji autokorelasi dapat menggunakan Durbin Watson Test (Gujarati, 1991). Jika hasil uji berada pada rentang nilai antara 2 (maksimal) dan 4-dl (minimal), maka disimpulkan bahwa tidak terjadi autokorelasi.

Uji heterokedastisitas

Uji heterokedastisitas perlu dilakukan untuk menganalisis apakah pada suatu model regresi terjadi ketidaknyamanan varian dari residual pada satu pengamatan terhadap pengamatan lainnya. Pada umumnya, data *crosssection* akan mengalami kondisi heterokedastisitas karena data ini menghimpun data yang mewakili berbagai ukuran: kecil, sedang, dan besar (Ghozali, 2016).

HASIL DAN PEMBAHASAN

Uji Multikolinearitas

Uji Multikolinearitas yang telah dilakukan menggunakan *software* Minitab 19 menghasilkan nilai sebagai berikut.

Coefficients

| Term | Coef | SE Coef | T-Value | P-Value | VIF |
|-------------------|-----------|----------|---------|---------|------|
| Constant | -1476 | 1205 | -1.22 | 0.288 | |
| PDRB (X1) | 0.008695 | 0.000926 | 9.39 | 0.001 | 1.80 |
| Jlh Penduduk (X2) | -0.000048 | 0.000120 | -0.40 | 0.710 | 2.09 |
| Inflasi (X3) | 646 | 604 | 1.07 | 0.345 | 1.26 |

Berdasarkan data

Correlations

| | Pajak Daerah (Y) | PDRB (X1) | Jlh Penduduk (X2) |
|-------------------|------------------|-----------|-------------------|
| PDRB (X1) | 0.984 | | |
| Jlh Penduduk (X2) | 0.654 | 0.665 | |
| Inflasi (X3) | 0.342 | 0.272 | 0.451 |

Gambar 1. Hasil Uji Multikolinearitas

Dari hasil tersebut diketahui bahwa terjadi multikolinearitas. Terlihat pada gambar 1, dari tiap variabel tidak menunjukkan nilai VIF > 10. Akan tetapi, terdapat perubahan tanda koefisien variabel X2 dengan nilai korelasi variabel X2 terhadap variabel Y, yaitu -0,000048 dan 0,654. Sehingga dapat disimpulkan bahwa terjadi kasus multikolinearitas.

Model Regresi Linier dengan Metode Backward Elimination

Metode Backward Elimination memiliki proses yaitu dengan memasukkan semua variabel bebas ke dalam model kemudian mengeliminasi satu per satu variabel bebas/independen dari model regresi yang terbentuk (Ekanayake, Rankothge, Weliwatta, & Jayasinghe, 2021). Berikut adalah proses pembentukan model dengan metode backward elimination.

Tabel 1. Eliminasi variabel bebas pada model

| Model | Variabel yang masuk Model | Variabel yang dieliminasi dari Model |
|-------|---|--------------------------------------|
| 1 | PDRB (X1), Jumlah Penduduk (X2), Inflasi (X3) | - |
| 2 | PDRB (X1), Inflasi (X3) | Jumlah Penduduk (X2) |
| 3 | PDRB (X1) | Inflasi (X3) |

Tabel 1 memperlihatkan ketika dengan metode backward elimination menggunakan 2 tahapan eliminasi variabel bebas/independen dari model regresi, dan diperoleh 3 model regresi. Model pertama melibatkan 3 variabel bebas (PDRB, Jumlah Penduduk, Inflasi). Model Kedua melibatkan 2 variabel bebas (PDRB, Inflasi), dengan mengeliminasi variabel Jumlah Penduduk. Model ketiga melibatkan 1 variabel bebas (PDRB), dengan mengeliminasi variabel Inflasi. Proses eliminasi variabel-variabel bebas tersebut akan dijelaskan lebih lanjut sebagai berikut.

| Model | Source | DF | Adj SS | Adj MS | F-Value | P-Value |
|-------|------------|----|---------|---------|---------|---------|
| 1 | Regression | 3 | 8455564 | 2818521 | 52.69 | 0.001 |
| | Error | 4 | 213953 | 53488 | | |
| | Total | 7 | 8669517 | | | |
| 2 | Regression | 2 | 8447007 | 4223504 | 94.91 | 0.000 |
| | Error | 5 | 222509 | 44502 | | |
| | Total | 7 | 8669517 | | | |
| 3 | Regression | 1 | 8394445 | 8394445 | 183.10 | 0.000 |
| | Error | 6 | 275071 | 45845 | | |
| | Total | 7 | 8669517 | | | |

Gambar 2. ANOVA

Gambar 2 menunjukkan hasil uji serentak dari model 1 sampai model 3, dengan penjelasan sebagai berikut:

1) Nilai $F_{hitung} = 52,69$. Sedangkan nilai dengan taraf nyata 0,05 diperoleh $F_{tabel} = F_{(3,7,0.05)} = 4,35$.

Karena $F_{hitung} > F_{tabel}$, maka disimpulkan bahwa secara serentak seluruh variabel bebas yang masuk pada model 1 memiliki pengaruh terhadap variabel terikat.

- 2) Nilai $F_{hitung} = 94,91$. Sedangkan nilai dengan taraf nyata 0,05 diperoleh $F_{tabel} = F_{(2,7,0.05)} = 4,74$.
 Karena $F_{hitung} > F_{tabel}$, maka disimpulkan bahwa secara serentak seluruh variabel bebas yang masuk pada model 2 memiliki pengaruh terhadap variabel terikat.
- 3) Nilai $F_{hitung} = 183,10$. Sedangkan nilai dengan taraf nyata 0,05 diperoleh $F_{tabel} = F_{(1,7,0.05)} = 5,59$.
 Karena $F_{hitung} > F_{tabel}$, maka disimpulkan bahwa secara serentak seluruh variabel bebas yang masuk pada model 3 memiliki pengaruh terhadap variabel terikat.

Karena model 1, model 2, dan model 3 memiliki hasil yang sama, maka akan dilanjutkan uji selanjutnya yaitu uji korelasi parsial.

| Model | Term | Coef | SE Coef | T-Value | P-Value |
|-------|-------------------|-----------|----------|---------|---------|
| 1 | Constant | -1476 | 1205 | -1.22 | 0.288 |
| | PDRB (X1) | 0.008695 | 0.000926 | 9.39 | 0.001 |
| | Jlh Penduduk (X2) | -0.000048 | 0.000120 | -0.40 | 0.710 |
| | Inflasi (X3) | 646 | 604 | 1.07 | 0.345 |
| 2 | Constant | -1327 | 1046 | -1.27 | 0.260 |
| | PDRB (X1) | 0.008461 | 0.000655 | 12.92 | 0.000 |
| | Inflasi (X3) | 555 | 510 | 1.09 | 0.327 |
| 3 | Constant | -194.7 | 92.9 | -2.09 | 0.081 |
| | PDRB (X1) | 0.008654 | 0.000640 | 13.53 | 0.000 |

Gambar 3. Hasil Uji Korelasi Parsial

Pengujian secara parsial pada ketiga model tersebut menunjukkan jika model 3 berpengaruh signifikan terhadap variabel terikat Y, karena setiap variabel bebas memiliki nilai signifikansi atau $p - value = 0.000 < \alpha = 0,05$. Sedangkan pada model 1 dan 2 masih terdapat variabel bebas yang nilai signifikansinya atau $p - value > \alpha = 0,05$.

Dari serangkaian proses yang dilakukan, dapat disimpulkan bahwa model yang dihasilkan dengan menggunakan metode backward elimination adalah model 3 dengan variabel bebas PDRB (X_1) dan variabel terikatnya Pajak Daerah (Y). Sehingga model regresi yang dibentuk adalah sebagai berikut:

$$Y = -194,7 + (0,008654)X_1$$

Dengan nilai koefisien determinasi sebesar 0,9683 atau 96,83% yang menunjukkan bahwa variabel X_1 berpengaruh sebesar 96,83% terhadap variabel Y, sedangkan sisanya dipengaruhi faktor lainnya.

Uji Asumsi Klasik Regresi Linier

Setelah diperoleh model regresi terbaik, akan dilakukan uji asumsi klasik berupa uji normalitas, uji heteroskedastisitas, dan uji autokorelasi.

Normalitas

Dengan Minitab 19 diperoleh hasil pengujian sebagai berikut

| n | Kolmogrov-Smirnov Z | P-Value |
|---|---------------------|---------|
| 8 | 0.649 | >0.150 |

Gambar 4. Hasil Uji Normalitas

Dari tabel tersebut diperoleh p-value atau signifikansi $> 0.150 > 0.05$, maka terima H_0 . Artinya residual menyebar berdistribusi normal.

Heteroskedastisitas

Dengan menggunakan software minitab 19, diperoleh hasil sebagai berikut

| Term | T-Value | P-Value |
|----------|---------|---------|
| Constant | 2.58 | 0.042 |
| PDRB | -0.21 | 0.841 |

Gambar 5. Hasil Uji Heteroskedastisitas

Dari tabel diatas diketahui bahwa nilai signifikansi atau p-value variabel PDRB sebesar 0.841. Karena nilai tersebut lebih besar dari α , maka terima H_0 sehingga disimpulkan bahwa tidak terjadi heteroskedastisitas.

Autokorelasi

Berikut hasil pengujian autokorelasi menggunakan *software* minitab 19

| R^2 | Durbin-Watson |
|--------|---------------|
| 0.9683 | 1.82908 |

Gambar 6. Hasil Uji Autokorelasi

Nilai Durbin-Watson yang diperoleh sebesar 1.82908. Sedangkan nilai tabel Durbin-Watson dengan $n=8$ dan $k=1$ adalah $d_L = 0.5591$ dan $d_U = 1.7771$. Karena nilai Durbin-Watson terletak antara d_U dan $4 - d_U$ ($1.7771 < 1.82908 < 2.2229$), maka terima H_0 yang artinya tidak ada autokorelasi residual.

PENUTUP

Berdasarkan rumusan masalah, pembahasan dan hasil penelitian model regresi terbaik dengan metode *backward elimination* pada regresi linier berganda, maka dapat diambil kesimpulan bahwa model regresi terbaik yang dihasilkan dengan metode *backward elimination* adalah $Y = -194,7 + (0,008654)X_1$, dan faktor yang paling berpengaruh pada pajak daerah dengan model regresi yang diperoleh adalah Produk Domestik Regional Bruto (PDRB) dengan koefisien determinasi sebesar 0,9683 atau 96,83%. Model regresi yang digunakan hanya diterapkan pada beberapa kabupaten/kota di Jawa Timur. Saran untuk penelitian selanjutnya adalah melakukan pemodelan menggunakan data yang lebih luas, yang dapat mewakili berbagai provinsi di seluruh wilayah Indonesia.

DAFTAR RUJUKAN

bps.go.id. (2021, Agustus 7).

djpk.kemenkeu.go.id. (2021, 8 5).

Ekanayake, P., Rankothge, W., Weliwatta, R., & Jayasinghe, J. W. (2021). Machine Learning Modelling of the Relationship between Weather and Paddy Yield in Sri Lanka. *Journal of Mathematics*, 1-14.

Ghozali, I. (2016). *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23* (8 ed.). Semarang: Badan Penerbit Universitas Diponegoro.

Gujarati, D. (1991). *Ekonometrika Dasar, Terjemahan*. Jakarta: Erlangga.

jatim.bps.go.id. (2021, 8 7).

Novelysa, S., Siagian, P., & Bangun, P. (2014). ANALISA METODE BACKWARD DAN METODE FORWARD UNTUK MENENTUKAN PERSAMAAN REGRESI LINIER BERGANDA (Studi Kasus: Jumlah Kecelakaan Lalu Lintas di Kotamadya Medan). *Saintia Matematika*, 2, 345-360.

Sulistiyono, & Sulistiyowati, W. (2017, Desember). Peramalan Produksi dengan Metode Regresi Linier Berganda. *PROZIMA*, 1, 82-89.

Yulianti, V. N., & Sembiring, P. (2023). Penerapan Metode Backward untuk Menentukan Persamaan Regresi Linier Berganda pada Dugaan Tindak Pidana di Kota Binjai. *FARABI : Jurnal Matematika dan Pendidikan Matematika*, 6, 1-9.