

**PENGARUH PENDIDIKAN, PENGALAMAN, UMUR, DAN JAM KERJA TERHADAP
PENDAPATAN BULANAN KARYAWAN MENGGUNAKAN MODEL REGRESI
LINEAR BERGANDA**

**Michael D Sianturi¹, Imelda Octavia L.Tobing², Fachriz Effendy. K³, Bintang Maulana⁴,
Gizka Triyunita Sinaga⁵, Yolanda Angelina Sitorus⁶**

Prodi Statistika, Fakultas Matematika, Universitas Negeri Medan

Email : imeldaoctaviatobing@gmail.com

ABSTRAK

Penelitian ini menganalisis pengaruh tingkat pendidikan, pengalaman kerja, umur, jenis kelamin, dan jumlah jam kerja terhadap pendapatan bulanan karyawan menggunakan model regresi linier berganda. Data dikumpulkan dari karyawan PT. Nubika Jaya Labuhan Batu Selatan. Hasil analisis menunjukkan bahwa umur dan jumlah jam kerja memiliki pengaruh positif signifikan terhadap pendapatan, sedangkan tingkat pendidikan dan pengalaman kerja menunjukkan pengaruh negatif. Model yang digunakan telah melalui uji asumsi klasik, termasuk uji normalitas, multikolinearitas, dan autokorelasi, dengan hasil bahwa asumsi terpenuhi. Dengan nilai Adjusted R-squared sebesar 30,69%, penelitian ini berhasil menjelaskan sebagian variasi pendapatan karyawan berdasarkan variabel independen yang digunakan. Temuan ini memberikan wawasan penting untuk pengembangan kebijakan ketenagakerjaan dan perencanaan karier individu.

Kata Kunci: Pendapatan Karyawan, Regresi Linier Berganda

ABSTRACT

This study analyzes the influence of education level, work experience, age, gender, and working hours on employees' monthly income using a multiple linear regression model. Data were collected from employees at PT. Nubika Jaya Labuhan Batu Selatan. The analysis results indicate that age and working hours positively and significantly affect income, while education level and work experience show negative effects. The model has undergone classical assumption tests, including normality, multicollinearity, and autocorrelation tests, all of which were satisfied. With an Adjusted R-squared value of 30.69%, this study successfully explains part of the variation in employees' income based on the independent variables used. These findings provide valuable insights for developing employment policies and individual career planning.

Keywords: Employee Income, Multiple Linear Regression

Article History

Received: Desember 2024

Reviewed: Desember 2024

Published: Desember 2024

Plagiarism Checker No 223

DOI : 10.8734/Trigo.v1i2.365

Copyright : Author

Publish by : Trigonometri



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

PENDAHULUAN

Pengaruh umur, pendidikan, pengalaman dan jam kerja terhadap pendapatan merupakan topik yang krusial dalam studi ekonomi tenaga kerja, karena keempat faktor ini sering kali memengaruhi potensi penghasilan individu. Secara umum, seiring bertambahnya usia, seseorang cenderung mengumpulkan lebih banyak pengalaman kerja, yang dapat meningkatkan keterampilan dan produktivitas. Pengalaman yang lebih lama biasanya disertai dengan pengetahuan yang lebih mendalam dan keahlian khusus dalam bidang tertentu, sehingga meningkatkan peluang untuk memperoleh posisi yang lebih tinggi dan pendapatan yang lebih besar.

Menurut (Mustofa & Pd, n.d.-a) daya produksi buruh mempunyai hubungan yang positif dengan taraf pendidikan dan latihan. Semakin tinggi taraf pendidikan dan latihan yang dimiliki oleh seseorang maka semakin produktif individu tersebut. Selanjutnya keadaan ini mewujudkan hubungan yang positif antara taraf pendidikan dengan pendapatan. Semakin tinggi pencapaian taraf pendidikan maka peningkatan daya pengeluaran, kemahiran, dan cara berfikir, dan kecakapan akan meningkatkan upah atau pendidikan. Hal ini di kemukakan oleh penelitian terdahulu bahwa Pendidikan berpengaruh pada produktifitas dan efisiensi kerja seseorang yang kemudian akan mempengaruhi Real Income individu atau rumah tangga. (Tri Aprilia, 2023).

Lalu pengalaman kerja adalah ukuran sebagaimana cepat atau lamanya waktu atau riwayat kerja seseorang yang telah ditempuh oleh karena itu akan dapat dengan mudah dan terlatih dalam memahami tugas – tugas dalam suatu bidang pekerjaannya. Semakin lama seseorang dalam bekerja sesuai dengan keahliannya diharapkan akan mampu meningkatkan produktivitasnya (Mustofa & Pd, n.d.-b). Dengan pengalaman kerja ini karyawan dapat dan mampu meningkatkan produktivitasnya di bidang pekerjaan yang sudah dipilih

Usia adalah batasan atau tingkat ukuran hidup yang mempengaruhi kondisi fisik seseorang. Biasanya individu yang berada di usia produktif (>30 tahun) memiliki literasi keuangan yang lebih tinggi dibanding dibawah ataupun di atas usia produktif. Umur juga merupakan variabel yang relevan, karena karyawan yang lebih tua biasanya memiliki pengalaman kerja yang lebih banyak dan memiliki posisi yang lebih stabil dalam pekerjaan mereka. Ini disebabkan karena pengetahuan yang dimiliki dan kemampuan berpikir secara logis. (Nurhayani, 2019)

Menurut (Mahendra et al., n.d.) Curahan jam kerja merupakan lama waktu yang digunakan seorang pekerja untuk melayani pekerjaannya, dalam penelitian ini dikaitkan dengan lamanya pengalaman karyawan dalam bekerja. Menurut (Agustinus Aryanto, 2017) curahan jam kerja adalah Jumlah kerja yang ditawarkan oleh tenaga kerja dengan menggunakan satuan jam kerja per minggu.

Oleh karena itu, analisis regresi linear berganda sering diterapkan untuk menilai pengaruh Pendidikan, Pengalaman, Umur, dan Jam Kerja secara bersamaan terhadap pendapatan, memungkinkan kita untuk mengidentifikasi kontribusi masing-masing faktor dan memperoleh pemahaman yang lebih mendalam mengenai dinamika pendapatan dalam berbagai sektor dan pekerjaan. Melalui analisis ini, kita dapat memprediksi pendapatan di masa depan berdasarkan variabel, yang dapat mendukung perencanaan karier individu serta pembuatan kebijakan ketenagakerjaan yang lebih tepat dan efektif.

Menurut (Sinaga et al., 2022) Regresi linier merupakan satu cara prediksi yang menggunakan garis lurus untuk menggambarkan hubungan diantara dua variabel atau lebih. Variabel adalah besaran yang berubah-ubah nilainya. Selanjutnya variabel tersebut terbagi atas

dua jenis yaitu variabel pemberi pengaruh dan variabel terpengaruh. Variabel pemberi pengaruh dapat dianalogikan sebab, sementara variabel terpengaruh merupakan akibat, Selanjutnya peramalan ini didasarkan pada asumsi bahwa pola pertumbuhan data historis yang bersifat linier, walaupun sebenarnya tidak 100% linier. Pola pertumbuhan ini didekati dengan suatu model yang menggambarkan hubungan-hubungan yang terkait dalam suatu keadaan. Regresi linier berganda merupakan model persamaan yang menjelaskan hubungan satu variabel tak bebas/ response (Y) dengan dua atau lebih variabel bebas/predictor (X_1, X_2, \dots, X_n). memprediksi nilai variabel tak bebas/ response (Y) apabila nilai-nilai variabel bebasnya/ predictor (X_1, X_2, \dots, X_n) diketahui. Disamping itu juga untuk dapat mengetahui bagaimanakah arah hubungan variabel tak bebas dengan variabel-variabel bebasnya.

Regresi berganda untuk menduga nilai dari parameter dan model yang dirumuskan sebagai berikut:

$$Y = X\beta + \varepsilon \quad (1)$$

Estimasi parameter $\beta_0, \beta_1, \beta_2, \beta_k$ diperoleh dengan meminimumkan jumlah kuadrat residual dari persamaan regresi. Residual persamaan (1) dinyatakan sebagai berikut:

$$\varepsilon = Y - \hat{Y} \quad (2)$$

Dengan metode kuadrat terkecil jumlah kuadrat residual dirumuskan sebagai berikut:

$$S(\beta) = y'y - 2\beta'X'y + \beta'X'X \quad (3)$$

Kemudian jumlah kuadrat residual di minimumkan melalui turunan parsial dan menyamakan persamaan dengan nol terhadap $\hat{\beta}$ yaitu:

$$\hat{\beta} = \frac{x'y}{x'x} \quad (4)$$

Dengan memastikan model regresi linier klasik estimator OLS memiliki beberapa sifat ideal atau optimum, dimana suatu estimator $\hat{\beta}$ dikatakan Best Linear Unbiased Estimator (BLUE) dari β jika estimator $\hat{\beta}$ linier terhadap Y tak bias dan mempunyai variansi yang minimum. Sehingga di peroleh (Prasetyo & Helma, 2022)

$$\hat{\beta} = \beta + \frac{x'\varepsilon}{x'x} \quad (5)$$

METODE PENELITIAN

1. Rancangan Penelitian

Penelitian ini merupakan penelitian kuantitatif yang menggunakan metode analisis regresi linear berganda untuk mengidentifikasi dan menganalisis pengaruh variabel independen terhadap pendapatan bulanan karyawan. Variabel independen yang digunakan meliputi tingkat pendidikan, pengalaman kerja, umur, jenis kelamin, dan jumlah jam kerja. Data penelitian dikumpulkan dari PT. Nubika Jaya Labuhan Batu Selatan. Data yang diperoleh selanjutnya dianalisis untuk mendapatkan pemahaman yang lebih mendalam mengenai hubungan antara variabel-variabel tersebut.

2. Identifikasi dan Pengukuran Variabel

Dalam penelitian ini, variabel yang diukur yaitu :

- Pendapatan Bulanan (Y) : Jumlah pendapatan yang diterima oleh karyawan dalam satu bulan (dalam satuan juta rupiah)
- Tingkat Pendidikan (X_1) : Tingkat pendidikan formal tertinggi yang dicapai oleh karyawan
- Pengalaman Kerja (X_2) : Lama waktu karyawan bekerja secara profesional dalam tahun
- Umur (X_3) : Usia karyawan saat ini

- e. Jenis Kelamin (X_4) : Jenis kelamin biologis karyawan
- f. Jumlah Jam Kerja (X_5) : Rata-rata jumlah jam kerja karyawan dalam satu minggu

3. Menguji Data

Pada tahap ini, dilakukan pengujian asumsi klasik dan pengujian hipotesis. Analisis ini dilakukan dengan membentuk model regresi linear berganda untuk memahami pengaruh variabel independen terhadap variabel dependen. Berikut adalah langkah-langkah yang telah dilakukan:

Model Regresi Linear Berganda

Model regresi dirumuskan sebagai:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon$$

Di mana:

Y : Pendapatan Bulanan (dalam satuan juta rupiah)

X_1 : Tingkat Pendidikan

X_2 : Pengalaman Kerja

X_3 : Umur

X_4 : Jenis Kelamin

X_5 : Jumlah Jam Kerja

Pengujian Hipotesis

Menggunakan uji t untuk menguji signifikansi masing-masing variabel independen dan uji F untuk menguji signifikansi model secara keseluruhan.

4. Pemeriksaan asumsi dilakukan untuk memastikan validitas model. Hasil yang diperoleh:

1. Uji Normalitas Residual

- o Dilakukan uji Shapiro-Wilk untuk mengevaluasi distribusi residual. Hasil menunjukkan residual berdistribusi normal, terlihat dari p-value > 0.05 pada uji Shapiro-Wilk dan visualisasi Q-Q plot yang mendekati garis diagonal.

2. Uji Multikolinearitas

- o Variance Inflation Factor (VIF) digunakan untuk mengevaluasi hubungan antar variabel independen. Semua nilai VIF berada di bawah ambang batas 10, menunjukkan tidak adanya multikolinearitas yang signifikan.

3. Uji Autokorelasi

- o Dilakukan uji Durbin-Watson, dan nilai DW mendekati 2, menunjukkan tidak adanya autokorelasi.

5. Melakukan pemilihan model terbaik

Pemilihan model terbaik dalam penelitian ini dilakukan dengan membandingkan beberapa spesifikasi model regresi linear berganda untuk menentukan model yang paling sesuai dengan data. Proses ini melibatkan evaluasi terhadap berbagai indikator kinerja model, seperti Adjusted R-squared, nilai standar error residual, dan signifikansi statistik koefisien variabel independen. Selain itu, uji asumsi klasik, seperti normalitas residual, multikolinearitas, dan autokorelasi, digunakan untuk memastikan bahwa model memenuhi asumsi dasar regresi linear. Model yang memenuhi asumsi ini dengan nilai Adjusted R-

squared yang tinggi serta koefisien yang signifikan pada sebagian besar variabel independen dipilih sebagai model terbaik.

HASIL DAN PEMBAHASAN

Pada hasil akhir dalam penelitian ini menyajikan proses pengolahan data menggunakan aplikasi R Studio dengan menggunakan metode Regresi Linier Berganda. Data yang digunakan adalah data dari

1. Pengolahan Data

Sampel data yang digunakan dalam menentukan pendapatan bulanan karyawan adalah sebagai berikut:

Tabel 1. Sampel Data

No	Tingkat Pendidikan	Pengalaman Kerja	Umur	Jenis Kelamin	Jumlah Jam kerja	Pendapatan Bulanan
1	1	8	35	0	45	6
2	2	30	54	1	48	5
3	2	25	48	0	48	5,5
4	1	24	48	1	48	7
5	1	10	34	0	45	6
6	1	13	37	0	38	4,5
7	1	24	48	0	40	6
8	1	14	52	1	42	6,5
9	1	10	32	1	42	6,5
10	1	15	36	1	42	6,5
11	3	6	35	1	42	5,5
12	2	4	31	1	42	5,5
13	3	5	27	1	42	5,3
14	3	12	33	1	42	5,5
15	1	12	35	1	42	5,5
16	3	5	28	1	42	5,2
17	1	5	27	1	42	5,3
18	3	4	31	1	42	5,2
19	3	2	27	1	42	5,2
20	2	1,5	34	1	42	5,2
21	2	12	39	1	42	5,3
22	3	10	31	1	42	5,4
23	1	5	38	1	42	6
24	2	12	35	1	42	5,1
25	2	10	34	1	42	5,7
26	3	4	27	1	42	5,2
27	2	11	37	1	42	5,8
28	1	10	40	1	42	6
29	3	3	25	1	42	5,4
30	1	11	42	1	42	5,6
31	3	6	29	1	42	5,3
32	3	8	35	1	42	5,4
33	3	5	31	1	42	5,3
34	2	13	39	1	42	5,5

35	2	15	42	1	42	5,6
36	1	7	44	1	42	5,8
37	2	10	47	1	42	5,4
38	1	7	38	1	42	5,8
39	3	11	32	1	42	5,3
40	3	9	36	1	42	5,4
41	2	3	28	1	42	5,3
42	1	10	43	1	42	5,1
43	3	3,5	25	1	42	5,2
44	1	3	25	1	40	4,5
45	2	12	32	0	42	4,5
46	3	20	40	1	45	6,5
47	2	5	27	1	38	4,5
48	1	8	30	0	43	4
49	3	15	50	0	46	6,5
50	1	3	22	1	40	4,5
51	2	10	35	1	41	5,2
52	3	18	45	0	47	6,5
53	2	7	28	1	44	4,5

Sumber: PT. Nubika Jaya Labuhan Batu Selatan

2. Deskripsikan Statistik Dasar

Setelah data set yang telah didapat dari responden, maka dilakukan deskripsi statistik dasar. Ini berguna untuk mengetahui rata – rata, median, standar deviasi, dll. Untuk mengetahui di R maka input code seperti pada Gambar 1.

```
> library(psych)
> describe(project_anreg)
      vars  n mean  sd median trimmed mad  min max range skew
Tingkat Pendidikan 1 50  1.84 0.87   2.0   1.80 1.48  1.0  3  2.0  0.30
Pengalaman Kerja   2 50 10.34 6.40  10.0   9.54 5.93  1.5 30 28.5  1.03
Umur                3 50 36.44 7.37  35.0  36.00 6.67 25.0 54 29.0  0.48
Jenis kelamin       4 50  0.86 0.35   1.0   0.95 0.00  0.0  1  1.0 -2.01
Jumlah Jam kerja    5 50 42.20 1.90  42.0  42.00 0.00 38.0 48 10.0  1.16
Pendapatan Bulanan 6 50  5.87 1.55   5.5   5.61 0.44  2.5 12  9.5  2.00
      kurtosis  se
Tingkat Pendidikan -1.62 0.12
Pengalaman Kerja   0.76 0.91
Umur               -0.63 1.04
Jenis kelamin       2.10 0.05
Jumlah Jam kerja    3.94 0.27
Pendapatan Bulanan 5.54 0.22
> |
```

Gambar 1. Deskripsi Statistik

Sumber: RStudio

Berikut ada interpretasi output yang ditampilkan

a. Tingkat Pendidikan

Pada Tingkat pendidikan cenderung terdistribusi antara kategori yang lebih rendah dan sedikit lebih banyak berada pada tingkat 2. Berdasarkan skewness yang positif, ada kecenderungan bahwa sebagian besar individu memiliki tingkat pendidikan yang lebih tinggi.

b. Pengalaman Kerja

Rata-rata pengalaman kerja adalah sekitar 10 tahun, namun ada variasi yang cukup

besar. Skewness positif menunjukkan sebagian besar individu memiliki pengalaman kerja yang lebih sedikit, sementara sedikit individu memiliki pengalaman yang sangat panjang.

c. Umur

Rata-rata umur responden adalah 36.44 tahun dengan sebagian besar berada di sekitar 35 tahun. Skewness yang positif menunjukkan bahwa sebagian besar individu lebih muda, tetapi ada beberapa individu yang lebih tua.

d. Jenis Kelamin

Data ini menunjukkan bahwa mayoritas responden adalah pria (karena nilai mean mendekati 1, yang mewakili pria), dengan skewness yang menunjukkan bahwa distribusinya sangat tidak seimbang (lebih banyak pria).

e. Jumlah Jam Kerja

Rata-rata jam kerja adalah 42.20 jam per minggu, dengan sedikit variasi di antara individu. Skewness positif menunjukkan bahwa sebagian besar individu bekerja sekitar 42 jam, namun ada sebagian kecil yang bekerja lebih lama.

f. Pendapatan Bulanan

Rata-rata pendapatan bulanan adalah sekitar 5.87 juta IDR, dengan sebagian besar pendapatan berada di sekitar angka tersebut. Skewness yang sangat tinggi menunjukkan ada sebagian kecil individu dengan pendapatan jauh lebih tinggi daripada rata-rata, sementara sebagian besar individu memiliki pendapatan yang lebih rendah. Kurtosis yang sangat tinggi juga menunjukkan distribusi pendapatan yang tidak merata, dengan lebih banyak individu memiliki pendapatan yang lebih rendah.

Oleh karena itu, Tingkat Pendidikan cenderung berada pada tingkat menengah (kategori 2), dengan sedikit bias ke arah yang lebih tinggi. Pengalaman Kerja dan Umur menunjukkan distribusi yang cukup merata dengan sedikit preferensi untuk kelompok yang lebih muda dan berpengalaman. Jenis Kelamin sangat tidak seimbang, dengan dominasi responden pria. Jumlah Jam Kerja menunjukkan bahwa sebagian besar individu bekerja sekitar 42 jam seminggu, dengan sedikit variasi. Pendapatan Bulanan sangat terpengaruh oleh sebagian kecil individu yang memiliki pendapatan jauh lebih tinggi, sementara sebagian besar memiliki pendapatan di bawah rata-rata. Analisis statistik ini memberi gambaran tentang distribusi dan karakteristik populasi dalam sampel, serta memberikan wawasan tentang faktor-faktor yang mempengaruhi pendapatan dan dinamika tenaga kerja.

3. Korelasi

Dengan menggunakan input yang di pakai di R di gambar 2 maka akan mendapat output sebagai berikut:

```
> # Tampilkan matriks korelasi
> print(correlation_matrix)
          Tingkat Pendidikan Pengalaman Kerja      Umur Jenis Kelamin Jumlah Jam kerja
Tingkat Pendidikan  1.00000000 -0.1558817 -0.2481674  0.1702612  0.1059788
Pengalaman Kerja   -0.15588168  1.0000000  0.7920653  -0.3557879  0.5765554
Umur                -0.24816737  0.7920653  1.0000000  -0.2595813  0.4983896
Jenis kelamin       0.17026118  -0.3557879 -0.2595813  1.0000000  -0.2974260
Jumlah Jam kerja    0.10597884  0.5765554  0.4983896  -0.2974260  1.0000000
Pendapatan_Bulanan -0.08655277  0.4141786  0.5411162  -0.0210659  0.4543553

          Pendapatan_Bulanan
Tingkat Pendidikan -0.08655277
Pengalaman Kerja   0.41417861
Umur                0.54111623
Jenis kelamin       -0.02106590
Jumlah Jam kerja    0.45435528
Pendapatan_Bulanan 1.00000000
```

Gambar 2. Korelasi Matrix

Sumber: RStudio

Dari output yang terdapat pada Gambar 2. Dapat diinterpretasikan sebagai berikut: Matriks ini menunjukkan korelasi antara beberapa variabel terkait dengan tingkat pendidikan, jenis kelamin, dan jumlah jam kerja pada suatu populasi. Berikut adalah interpretasi dari angka-angka yang ditampilkan:

- a. Tingkat Pendidikan memiliki korelasi positif dengan Pengalaman Kerja (0.1702612) dan Jumlah Jam Kerja (0.1059788), menunjukkan bahwa semakin tinggi tingkat pendidikan, semakin banyak pengalaman kerja dan jam kerja.
- b. Jenis Kelamin memiliki korelasi negatif dengan Tingkat Pendidikan (-0.1558168), Pengalaman Kerja (-0.2955879), dan Jumlah Jam Kerja (-0.2974260), mengindikasikan bahwa jenis kelamin tertentu cenderung memiliki tingkat pendidikan, pengalaman kerja, dan jam kerja yang lebih rendah.
- c. Umur memiliki korelasi negatif dengan Tingkat Pendidikan (-0.2481674) dan Jumlah Jam Kerja (-0.4983896), menunjukkan bahwa semakin tua usia, tingkat pendidikan dan jam kerja cenderung menurun.
- d. Jenis Kelamin memiliki korelasi positif dengan Pendapatan/Bulanan (0.4414786), mengindikasikan bahwa jenis kelamin tertentu memiliki pendapatan bulanan yang lebih tinggi.
- e. Umur memiliki korelasi negatif dengan Pendapatan/Bulanan (-0.0210659), menunjukkan bahwa semakin tua usia, pendapatan bulanan cenderung menurun.

Secara keseluruhan, matriks ini menggambarkan adanya beberapa hubungan menarik antara karakteristik demografis (tingkat pendidikan, jenis kelamin, umur) dan variabel terkait seperti pengalaman kerja, jam kerja, dan pendapatan bulanan pada populasi yang dianalisis.

4. Regresi Linier Berganda

```
> # Ringkasan hasil model
> summary(model)

Call:
lm(formula = `Pendapatan Bulanan` ~ `Tingkat Pendidikan` + `Pengalaman Kerja` +
    umur + `Jenis kelamin` + `Jumlah Jam kerja`, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-1.57407 -0.28989  0.03131  0.28968  1.13985

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -0.21797    1.81858  -0.120  0.9051
`Tingkat Pendidikan` -0.03147    0.09160  -0.344  0.7327
`Pengalaman kerja` -0.01226    0.02025  -0.605  0.5479
Umur           0.04082    0.01568   2.603  0.0123 *
`Jenis kelamin`   0.27818    0.20485   1.358  0.1810
`Jumlah Jam kerja` 0.09876    0.04501   2.194  0.0332 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5107 on 47 degrees of freedom
Multiple R-squared:  0.3735,    Adjusted R-squared:  0.3069
F-statistic: 5.605 on 5 and 47 DF,  p-value: 0.0003945
```

Gambar 3. Regresi Linier Berganda

Sumber: RStudio

Berikut adalah interpretasi hasil dari model regresi linear berganda yang ditampilkan dalam Gambar 3

a. Informasi Model:

- Persamaan model regresi: $\text{Pendapatan Bulanan} = -0.21797 + (-0.03147) * \text{Tingkat Pendidikan} + (-0.01226) * \text{Pengalaman Kerja} + 0.04082 * \text{Umur} + 0.27818 * \text{Jenis Kelamin} + 0.09876 * \text{Jumlah Jam Kerja}$

b. Residuals:

- Min: -1.57407
- Median: 0.03131
- 3Q: 0.28969
- Max: 1.13985
- Nilai residual menunjukkan perbedaan antara nilai observasi dan nilai prediksi dari model.

c. Koefisien:

- (Intercept): -0.21797 Ini adalah nilai konstanta atau intersep dari model regresi.
- Tingkat Pendidikan: -0.03147. Koefisien negatif menunjukkan bahwa setiap peningkatan satu unit tingkat pendidikan, pendapatan bulanan akan menurun sebesar 0.03147 unit, ketika variabel lain konstan.
- Pengalaman Kerja: -0.01226. Koefisien negatif menunjukkan bahwa setiap peningkatan satu unit pengalaman kerja, pendapatan bulanan akan menurun sebesar 0.01226 unit, ketika variabel lain konstan.
- Umur: 0.04082. Koefisien positif menunjukkan bahwa setiap peningkatan satu unit umur, pendapatan bulanan akan meningkat sebesar 0.04082 unit, ketika variabel lain konstan.
- Jenis Kelamin: 0.27812. Koefisien positif menunjukkan bahwa jenis kelamin tertentu memiliki pendapatan bulanan yang lebih tinggi sebesar 0.27818 unit, ketika variabel lain konstan.
- Jumlah Jam Kerja: 0.09876. Koefisien positif menunjukkan bahwa setiap peningkatan satu unit jumlah jam kerja, pendapatan bulanan akan meningkat sebesar 0.09876 unit, ketika variabel lain konstan.

d. Signifikansi:

- Tanda bintang (*) menunjukkan signifikansi statistik. Umur memiliki signifikansi pada level 0.05.

e. Asumsi Model:

- Residual standard error: 0.5107 pada 47 derajat kebebasan
- Multiple R-squared: 0.3735, Adjusted R-squared: 0.3069
- F-statistic: 5.605 pada 5 dan 47 DF, p-value: 0.0003945

f. Uji Simultan:

Uji simultan, atau yang juga disebut uji F, digunakan untuk mengetahui apakah semua variabel independen secara bersama-sama (simultan) berpengaruh signifikan terhadap variabel dependen. Dalam model ini, uji simultan ditunjukkan oleh nilai F-statistic yang bernilai 5.605 dengan p-value 0.0003945. Ini berarti bahwa setidaknya satu dari variabel independen (Tingkat Pendidikan, Pengalaman Kerja, Umur, Jenis Kelamin, Jumlah Jam Kerja) secara signifikan mempengaruhi variabel dependen (Pendapatan Bulanan).

g. Uji Parsial:

Uji parsial, atau yang juga disebut uji t, digunakan untuk mengetahui pengaruh masing-masing variabel independen secara individual terhadap variabel dependen. Dalam model ini, uji parsial ditunjukkan oleh nilai t-value dan p-value untuk masing-masing koefisien. Variabel Umur dengan p-value 0.0123 menunjukkan bahwa variabel Umur secara parsial berpengaruh signifikan terhadap Pendapatan Bulanan pada tingkat signifikansi 5%. Sedangkan untuk variabel lainnya, tidak menunjukkan pengaruh yang signifikan secara parsial.

h. Koefisien Determinasi:

Koefisien determinasi (R-squared) menunjukkan seberapa besar proporsi variasi dalam variabel dependen yang dapat dijelaskan oleh variasi dalam variabel-variabel independen. Dalam model ini, nilai R-squared adalah 0.3735, yang berarti bahwa 37,35% variasi dalam Pendapatan Bulanan dapat dijelaskan oleh variasi dalam Tingkat Pendidikan, Pengalaman Kerja, Umur, Jenis Kelamin, dan Jumlah Jam Kerja. Nilai Adjusted R-squared yang lebih rendah, yaitu 0.3069, menunjukkan bahwa model ini lebih baik dalam menjelaskan data sampel daripada populasi.

5. Uji Asumsi Klasik

a. Normalitas

```
> shapiro.test(residuals(model))  
  
shapiro-wilk normality test  
  
data: residuals(model)  
W = 0.96513, p-value = 0.1237
```

Gambar 4. Uji Normalitas

Sumber: RStudio

Uji Normalitas bertujuan untuk menguji apakah dalam model regresi variabel pengganggu atau residual memiliki distribusi normal(Frischa et al., 2021.)

Hasil uji Normalitas untuk data tersebut dapat dilihat pada gambar 4. Berdasarkan output yang diberikan didapatkan nilai signifikansi sebesar 0.1237. lebih dari alpha (0.05) maka dapat dikatakan bahwa galat berdistribusi normal.

b. Uji Multikolinearitas

```
> print(vif_values)  
`Tingkat Pendidikan`  `Pengalaman Kerja`  Umur  
1.189670             3.172266             2.884146  
`Jenis Kelamin`     `Jumlah Jam kerja`  
1.202179             1.671338
```

Gambar 5. Uji Mutikolinearitas

Sumber: RStudio

Berdasarkan hasil output pada gambar 5, didapatkan nilai VIF yang dimana variable variable ini dapat dikatakn ragam homogen karena lebih dari alpha (0.05). Nilai VIF < 10 maka artinya tidak terjadi multikolinieritas dalam model regresi.

c. Uji Autokorelasi

```
> dwtest(model)

Durbin-watson test

data: model
Dw = 1.8525, p-value = 0.2628
alternative hypothesis: true autocorrelation is greater than 0
```

Gambar 6. Uji Mutikolinearitas

Sumber: RStudio

Berdasarkan hasil diatas, didapatkan nilai signifikansi sebesar 0.2628 lebih dari alpha (0.05) maka dapat dikatakan bahwa asumsi non autokorelasi terpenuhi.

6. Deteksi Pencilan

```
> # Hitung Z-score
> z_scores <- scale(data$`Pendapatan Bulanan`)
> # Identifikasi pencilan
> pencilan_z <- which(abs(z_scores) > 3)
> pencilan_z
integer(0)
```

Gambar 7. Uji Mutikolinearitas

Sumber: RStudio

Kode R yang diberikan berfungsi untuk mengidentifikasi data pencilan (outlier) dalam variabel "Pendapatan Bulanan". Pertama, kode menghitung z-score untuk setiap data titik dalam variabel tersebut. Z-score merupakan ukuran standar deviasi dari rata-rata, sehingga nilai z-score yang sangat besar (biasanya di atas 3 atau di bawah -3) mengindikasikan bahwa data titik tersebut jauh berbeda dari data lainnya dan dapat dianggap sebagai pencilan. Kemudian, kode mencari data mana saja yang memiliki nilai z-score yang melebihi ambang batas tersebut. Hasil akhirnya adalah sebuah vektor yang berisi indeks dari data yang dianggap sebagai pencilan. Jika tidak ada data yang memenuhi kriteria pencilan, maka hasilnya akan berupa vektor kosong.

7. Interpretasi Koefisien

```
> print(interpretasi_koefisien)
# A tibble: 6 x 6
  term                estimate std.error statistic  p.value Interpretasi
  <chr>                <dbl>    <dbl>    <dbl>    <dbl> <chr>
1 (Intercept)         4.87     2.18      2.23 3.20e- 2 Ketika semua variabel...
2 `Tingkat Pendidikan` -0.303   0.0273  -11.1 2.54e-13 Setiap kenaikan satu ...
3 `Pengalaman kerja`  0.0108   0.00402  2.70 1.05e- 2 Setiap tambahan satu ...
4 Umur                -0.00434 0.00385  -1.13 2.67e- 1 Setiap tambahan satu ...
5 `Jenis kelamin`     0.107    0.212    0.506 6.16e- 1 Jika individu memilik...
6 `Jumlah Jam kerja`  0.0296   0.0481   0.614 5.43e- 1 Setiap tambahan satu ...
```

Gambar 8 . Interpretasi Koefisien

Sumber: RStudio

a. Intercept: Nilai: 4.87

Ketika semua variabel independen bernilai nol (atau berada pada nilai rata-rata jika variabel tersebut sudah distandarisasi), nilai prediksi untuk variabel dependen adalah 4.87. Namun, interpretasi intercept seringkali kurang relevan secara praktis, terutama jika nilai nol untuk variabel independen tidak masuk akal dalam konteks data.

b. Tingkat Pendidikan: Nilai: -0.303

Setiap kenaikan satu satuan pada tingkat pendidikan akan menyebabkan penurunan nilai variabel dependen sebesar 0.303 unit, dengan asumsi variabel independen lainnya konstan. Tanda minus menunjukkan hubungan negatif antara tingkat pendidikan dan variabel dependen.

c. Pengalaman Kerja: Nilai: 0.0108

Setiap tambahan satu tahun pengalaman kerja akan meningkatkan nilai variabel dependen sebesar 0.0108 unit, dengan asumsi variabel independen lainnya konstan.

d. Umur: Nilai: -0.0034

Interpretasi: Setiap penambahan satu tahun usia akan menyebabkan penurunan nilai variabel dependen sebesar 0.0034 unit, dengan asumsi variabel independen lainnya konstan.

e. Jenis Kelamin: Nilai: 0.107

Jika individu memiliki karakteristik yang diwakili oleh variabel "Jenis Kelamin" (misalnya, jika Jenis Kelamin dikodekan sebagai 1 untuk laki-laki dan 0 untuk perempuan), maka nilai variabel dependen akan meningkat sebesar 0.107 unit, dibandingkan dengan individu yang tidak memiliki karakteristik tersebut, dengan asumsi variabel independen lainnya konstan.

f. Jumlah Jam Kerja: Nilai: 0.0296

Setiap tambahan satu jam kerja akan meningkatkan nilai variabel dependen sebesar 0.0296 unit, dengan asumsi variabel independen lainnya konstan.

g. Signifikansi Statistik:

Perhatikan nilai p-value untuk setiap koefisien. Jika nilai p-value kurang dari 0.05, maka koefisien tersebut dianggap signifikan secara statistik, artinya pengaruh variabel independen terhadap variabel dependen tidak terjadi secara kebetulan.

h. Asumsi Regresi

Interpretasi pada gambar 8 didasarkan pada asumsi-asumsi regresi linear, seperti linearitas, tidak adanya multikolinearitas, homoskedastisitas, dan normalitas residual. Jika asumsi-asumsi ini tidak terpenuhi, maka interpretasi koefisien mungkin tidak valid.

8. Prediksi

```
# Rename columns in the original dataset (optional)
names(data) <- c("Pendidikan", "Pengalaman", "Umur", "Jeniskelamin", "JamKerja", "Pendapatan")

# Buat model dengan nama kolom yang baru
model <- lm(Pendapatan ~ Pendidikan + Pengalaman + Umur + Jeniskelamin + JamKerja, data = data)

# Data baru untuk prediksi (tanpa spasi di kolom) dengan 10 orang
new_data <- data.frame(
  Pendidikan = c(1, 2, 1, 3, 2, 1, 2, 3, 1, 2),
  Pengalaman = c(5, 15, 7, 12, 18, 10, 25, 8, 20, 14),
  Umur = c(30, 45, 35, 28, 50, 32, 38, 27, 40, 42),
  Jeniskelamin = c(0, 1, 1, 0, 0, 1, 1, 0, 0, 1),
  JamKerja = c(40, 50, 38, 42, 48, 45, 47, 35, 44, 50)
)

# Prediksi dengan model
predictions <- predict(model, newdata = new_data)

# Gabungkan data baru dengan hasil prediksi
results <- cbind(new_data, Pendapatan_Prediksi = predictions)

# Tampilkan hasil prediksi
print(results)
```

Gambar 9 . Prediksi

Sumber: RStudio

Kode R pada gambar 9 melakukan analisis regresi linear untuk memprediksi nilai "Pendapatan" berdasarkan beberapa faktor seperti "Pendidikan", "Pengalaman Kerja", "Umur", "Jenis Kelamin", dan "Jumlah Jam Kerja". Pertama, kode mempersiapkan data dengan memberikan nama yang jelas pada setiap kolom. Kemudian, model regresi dibangun untuk menemukan hubungan antara variabel-variabel tersebut. Setelah model terbentuk, kode menggunakan model ini untuk memprediksi nilai "Pendapatan" pada data baru. Hasil prediksi kemudian dibandingkan dengan nilai aktual untuk mengevaluasi kinerja model. Proses ini melibatkan langkah-langkah seperti perhitungan metrik kesalahan prediksi dan visualisasi data. Namun, untuk mendapatkan hasil yang lebih akurat dan reliabel, perlu dilakukan analisis lebih lanjut, seperti mengecek asumsi-asumsi regresi dan melakukan penyetulan model jika diperlukan.

Secara sederhana, kode ini membantu kita memahami bagaimana faktor-faktor tertentu dapat mempengaruhi pendapatan seseorang dan dapat digunakan untuk membuat prediksi tentang pendapatan di masa depan.

	Pendidikan	Pengalaman	Umur	JenisKelamin	JamKerja	Pendapatan_Prediksi
1	1	5	30	0	40	4.864103
2	2	15	45	1	50	6.588086
3	1	7	35	1	38	5.124349
4	3	12	28	0	42	4.831235
5	2	18	50	0	48	6.279705
6	1	10	32	1	45	5.656423
7	2	25	38	1	47	5.883513
8	3	8	27	0	35	4.148146
9	1	20	40	0	44	5.483451
10	2	14	42	1	50	6.477889

Gambar 10 . Input Prediksi

Sumber: RStudio

Menyajikan hasil prediksi pendapatan berdasarkan beberapa faktor seperti tingkat pendidikan, pengalaman kerja, usia, jenis kelamin, dan jumlah jam kerja. Setiap baris dalam tabel mewakili satu individu, di mana kolom "Pendapatan_Prediksi" menunjukkan pendapatan yang diprediksi oleh model untuk individu tersebut berdasarkan nilai-nilai variabel independennya. Dengan kata lain, model telah berusaha memperkirakan berapa pendapatan seseorang berdasarkan karakteristik yang dimiliki. Namun, penting untuk diingat bahwa prediksi ini tidak selalu sempurna dan bisa saja terdapat perbedaan antara nilai prediksi dengan nilai aktual. Perbedaan ini bisa disebabkan oleh berbagai faktor, seperti keterbatasan model, adanya variabel penting yang belum dimasukkan, atau adanya kesalahan dalam data. Untuk mendapatkan pemahaman yang lebih baik tentang akurasi model dan pengaruh masing-masing variabel, perlu dilakukan analisis lebih lanjut, seperti menghitung metrik evaluasi, membuat visualisasi data, dan menguji signifikansi koefisien regresi.

KESIMPULAN

Penelitian ini menyimpulkan bahwa variabel umur dan jumlah jam kerja memiliki pengaruh positif signifikan terhadap pendapatan bulanan karyawan, sedangkan tingkat pendidikan dan pengalaman kerja menunjukkan pengaruh negatif. Variabel jenis kelamin juga memberikan pengaruh positif, meskipun tidak signifikan secara statistik dalam model ini. Model regresi linier berganda yang digunakan memenuhi asumsi klasik, termasuk uji normalitas,

multikolinearitas, dan autokorelasi. Nilai Adjusted R-squared sebesar 30,69% menunjukkan bahwa variabel independen yang dianalisis mampu menjelaskan sebagian besar variasi pendapatan bulanan karyawan.

Temuan ini memberikan wawasan berharga bagi perumusan kebijakan ketenagakerjaan yang lebih tepat, khususnya dalam pengelolaan faktor-faktor yang memengaruhi pendapatan karyawan. Selain itu, hasil penelitian ini dapat menjadi dasar untuk penelitian lanjutan guna meningkatkan akurasi model atau menambahkan variabel lain yang relevan untuk meningkatkan pemahaman tentang dinamika pendapatan.

DAFTAR PUSTAKA

- Agustinus Aryanto, D. (2017). ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI PENDAPATAN BURUH NELAYAN
- Aprilia, T. (2023). *PENGARUH PENDIDIKAN, PENGALAMAN KERJA DAN PENDAPATAN TERHADAP LITERASI KEUANGAN IBU RUMAH TANGGA DI KECAMATAN MASAMBA KABUPATEN LUWU UTARA* (Doctoral dissertation, UNIVERSITAS MUHAMMADIYAH PALOPO).
- Frischa, *, Dwinantarie, P., Balafif, M., & Rahmasari, A. (2022). *Volume 2 Nomor 2 2021 Analisis Pengaruh Pengalaman Kerja, Usia, Jam Kerja dan Tingkat Pendidikan terhadap Pendapatan Karyawan di Departement Store Sogo Pakuwon Mall Surabaya*. <https://doi.org/10.46821/bharanomics.v2i2.258>
- Putra, P. M. S., & Kartika, N. (2019). Analisis pengaruh modal, umur, jam kerja, pengalaman kerja, dan pendidikan terhadap pendapatan nelayan di Kedonganan. *Jurnal Ekonomi Pembangunan Universitas Udayana*, 2(2), 1-20.
- Padilah, T. N., & Adam, R. I. (2019). Analisis regresi linier berganda dalam estimasi produktivitas tanaman padi di Kabupaten Karawang. *FIBONACCI: Jurnal Pendidikan Matematika Dan Matematika*, 5(2), 117-128.
- Mustofa, S., & Pd. (n.d.-a). *PENGARUH lryNrtmil(Ani, JIII(LAa IaM rruR.ra, DAII PENGALAMATI KERJA TERHADAP PENDAPATAI{ TENAGA KERTA II\NJUT USIA DI II\IIX)NE,SIA*.
- Nurhayati.(2019). *PENGARUH USIA, PENDIDIKAN, DAN PENDAPATAN TERHADAP LITERASI KEUANGAN PADA IBU RUMAH TANGGA (STUDI KASUS KELURAHAN BAGAN BATU KECAMATAN BAGAN SINEMBAH KABUPATEN ROKAN HILIR)*
- Prasetyo, R. A., & Helma, H. (2022). Analisis Regresi Linear Berganda Untuk Melihat Faktor Yang Berpengaruh Terhadap Kemiskinan di Provinsi Sumatera Barat. *Journal of Mathematics UNP*, 7(2), 62-68.
- Sinaga, W. A. L., Sumarno, S., & Sari, I. P. (2022). The Application of Multiple Linear Regression Method for Population Estimation Gunung Malela District. *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, 1(1), 55–64. <https://doi.org/10.55123/jomlai.v1i1.143>