

Tugas Laporan Analisis Regresi Linier Berganda Data Penjualan

Nama : Risiko

NIM : 241012000171

A. Latar Belakang dan Tujuan

Latar Belakang

Pendapatan harian merupakan indikator utama kinerja bisnis ritel. Banyak faktor yang memengaruhi pendapatan, antara lain jumlah pelanggan, pengeluaran iklan, serta tingkat diskon yang diberikan. Pemahaman hubungan antara faktor-faktor ini sangat penting bagi manajemen untuk merumuskan strategi yang tepat. Oleh karena itu, dilakukan analisis regresi linier berganda untuk mengetahui pengaruh jumlah pelanggan, pengeluaran iklan, dan tingkat diskon terhadap pendapatan harian.

Tujuan

- Menganalisis pengaruh jumlah pelanggan, pengeluaran iklan, dan tingkat diskon terhadap pendapatan harian.
- Menguji signifikansi variabel-variabel independen secara parsial dan simultan.
- Melakukan uji asumsi klasik sebagai syarat validitas model regresi.
- Memberikan rekomendasi berdasarkan hasil analisis.

B. Hasil dan Analisis

a. Statistik Deskriptif

| Variabel | Mean | Std | Min | 25% | Median | 75% | Max |
|-------------------|---------|---------|---------|---------|---------|---------|---------|
| Pendapatan_Harian | 5398.22 | 2029.56 | 1181.71 | 3852.96 | 5601.80 | 7082.44 | 9191.73 |
| Jumlah_Pelanggan | 178.53 | 73.71 | 43.00 | 116.75 | 181.50 | 239.75 | 299.00 |
| Pengeluaran_Ikl | 843.72 | 401.17 | 101.84 | 550.72 | 858.57 | 1202.74 | 1492.28 |

an

| | | | | | | | |
|----------|-------|------|------|------|-------|-------|-------|
| Tingkat_ | 14.77 | 8.80 | 0.09 | 6.82 | 14.58 | 21.44 | 29.87 |
| Diskon | | | | | | | |

Berikut *capture* menggunakan Python.

```
regression_analysis.py > ...
1
2 import pandas as pd
3 import statsmodels.api as sm
4 from scipy.stats import shapiro
5 from statsmodels.stats.outliers_influence import variance_inflation_factor
6 from statsmodels.stats.diagnostic import het_breuschpagan
7 import datetime
8
9 # 1. Membaca Dataset
10 df = pd.read_csv('data_penjualan.csv')
11
12 # 2. Statistik Deskriptif
13 desc = df.describe().T
14 print(f"Statistik Deskriptif: \n {desc}\n")
15
```

PROBLEMS OUTPUT DEBUG CONSOLE **TERMINAL** PORTS COMMENTS

`/media/r/d/K/Semester_2/STATISTIC AND DATA ANALYSIS /Tugas/TM-11`
python regression_analysis.py

Statistik Deskriptif:

| | count | mean | std | min | 25% | 50% | 75% | max |
|-------------------|-------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Pendapatan Harian | 200.0 | 5398.223841 | 2029.555259 | 1181.713514 | 3852.955098 | 5601.803654 | 7082.440165 | 9191.725587 |
| Jumlah Pelanggan | 200.0 | 178.525000 | 73.705781 | 43.000000 | 116.750000 | 181.500000 | 239.750000 | 299.000000 |
| Pengeluaran Iklan | 200.0 | 843.723308 | 401.174208 | 101.839806 | 550.717401 | 858.573972 | 1202.744722 | 1492.279696 |
| Tingkat_Diskon | 200.0 | 14.767720 | 8.804099 | 0.094183 | 6.824511 | 14.580946 | 21.441724 | 29.869259 |

b. Hasil Model Regresi

Model Regresi:

Pendapatan Harian = $-23.35 + 27.28 * \text{Jumlah Pelanggan} + 0.87 * \text{Pengeluaran Iklan} - 12.43 * \text{Tingkat Diskon}$

R-squared : 0.983

F-statistic : 3773 ($p < 0.001$)

Tabel Koefisien Regresi:

Intercept : -23.35 ($p=0.753$)

Jumlah Pelanggan : 27.28 ($p=0.000$)

Pengeluaran Iklan : 0.87 ($p=0.000$)

Tingkat Diskon : -12.43 ($p=0.000$)

Berikut *capture* Menggunakan Python :

```
regression_analysis.py > ...
9 # 1. Membaca Dataset
10 df = pd.read_csv('data_penjualan.csv')
11
12 # 2. Statistik Deskriptif
13 desc = df.describe().T
14 print(f"Statistik Deskriptif: \n {desc}\n")
15
16
17 # 3. Model Regresi Linier Berganda
18 X = df[['Jumlah_Pelanggan', 'Pengeluaran_Iklan', 'Tingkat_Diskon']]
19 y = df['Pendapatan_Harian']
20 X_const = sm.add_constant(X)
21 model = sm.OLS(y, X_const).fit()
22 reg_summary = model.summary()
23
24 print(f"Ringkasan Model Regresi: \n {reg_summary}\n")
25
26 # 4. Uji t, Uji F, Interpretasi Koefisien
27 params = model.params
28 pvalues = model.pvalues
29
30 print("Koefisien Model:")
31 for feature, coef, pval in zip(X.columns, params[1:], pvalues[1:]):
32     print(f"{feature}: Koefisien = {coef:.4f}, p-value = {pval:.4f}")
33
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS

Ringkasan Model Regresi:

OLS Regression Results

| | | | |
|-------------------|-------------------|---------------------|-----------|
| Dep. Variable: | Pendapatan_Harian | R-squared: | 0.983 |
| Model: | OLS | Adj. R-squared: | 0.983 |
| Method: | Least Squares | F-statistic: | 3773. |
| Date: | Fri, 06 Jun 2025 | Prob (F-statistic): | 4.77e-173 |
| Time: | 20:02:08 | Log-Likelihood: | -1399.1 |
| No. Observations: | 200 | AIC: | 2806. |
| Df Residuals: | 196 | BIC: | 2819. |
| Df Model: | 3 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P> t | [0.025 | 0.975] |
|-------------------|----------|---------|---------|-------|----------|---------|
| const | -23.3486 | 73.973 | -0.316 | 0.753 | -169.234 | 122.537 |
| Jumlah_Pelanggan | 27.2822 | 0.258 | 105.877 | 0.000 | 26.774 | 27.790 |
| Pengeluaran_Iklan | 0.8707 | 0.047 | 18.386 | 0.000 | 0.777 | 0.964 |
| Tingkat_Diskon | -12.4337 | 2.149 | -5.786 | 0.000 | -16.671 | -8.196 |

| | | | |
|----------------|-------|-------------------|----------|
| Omnibus: | 0.632 | Durbin-Watson: | 1.986 |
| Prob(Omnibus): | 0.729 | Jarque-Bera (JB): | 0.652 |
| Skew: | 0.134 | Prob(JB): | 0.722 |
| Kurtosis: | 2.917 | Cond. No. | 3.72e+03 |

c. Hasil Uji t dan Uji F

- Uji t: Seluruh variabel independen signifikan ($p < 0.05$).
- Uji F: Model secara keseluruhan signifikan ($\text{Prob}(F) \approx 0.000$).

Berikut capture menggunakan Python.

```
regression_analysis.py > ...
9 # 1. Membaca Dataset
10 df = pd.read_csv('data_penjualan.csv')
11
12 # 2. Statistik Deskriptif
13 desc = df.describe().T
14 print(f"Statistik Deskriptif: \n {desc}\n")
15
16
17 # 3. Model Regresi Linier Berganda
18 X = df[['Jumlah_Pelanggan', 'Pengeluaran_Iklan', 'Tingkat_Diskon']]
19 y = df['Pendapatan_Harian']
20 X_const = sm.add_constant(X)
21 model = sm.OLS(y, X_const).fit()
22 reg_summary = model.summary()
23
24 print(f"Ringkasan Model Regresi: \n {reg_summary}\n")
25
26 # 4. Uji t, Uji F, Interpretasi Koefisien
27 params = model.params
28 pvalues = model.pvalues
29
30 print("Koefisien Model:")
31 for feature, coef, pval in zip(X.columns, params[1:], pvalues[1:]):
32     print(f"{feature}: Koefisien = {coef:.4f}, p-value = {pval:.4f}")
33 # Interpretasi Uji t
34 for feature, pval in zip(X.columns, pvalues[1:]):
35     if pval < 0.05:
36         print(f"{feature} berpengaruh signifikan terhadap Pendapatan Harian (p-value = {pval:.4f})")
37     else:
38         print(f"{feature} tidak berpengaruh signifikan terhadap Pendapatan Harian (p-value = {pval:.4f})")
39 # Interpretasi Uji F
40 if model.f_pvalue < 0.05:
41     print(f"Model secara keseluruhan signifikan (p-value = {model.f_pvalue:.4f})")
42
43
44 # 5. Uji Normalitas Residual
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS COMMENTS

```
Koefisien Model:
Jumlah Pelanggan: Koefisien = 27.2822, p-value = 0.0000
Pengeluaran Iklan: Koefisien = 0.8707, p-value = 0.0000
Tingkat Diskon: Koefisien = -12.4337, p-value = 0.0000
Jumlah Pelanggan berpengaruh signifikan terhadap Pendapatan Harian (p-value = 0.0000)
Pengeluaran Iklan berpengaruh signifikan terhadap Pendapatan Harian (p-value = 0.0000)
Tingkat Diskon berpengaruh signifikan terhadap Pendapatan Harian (p-value = 0.0000)
Model secara keseluruhan signifikan (p-value = 0.0000)
```

d. Interpretasi Koefisien

- Jumlah Pelanggan: Setiap kenaikan 1 pelanggan, pendapatan harian bertambah rata-rata Rp27.280, signifikan dan hubungan positif.
- Pengeluaran Iklan: Setiap kenaikan Rp1 pada iklan, pendapatan harian naik Rp870, signifikan dan hubungan positif.
- Tingkat Diskon: Setiap kenaikan 1% diskon, pendapatan harian turun Rp12.430, signifikan dan hubungan negatif.

e. Hasil Uji Asumsi Klasik

Normalitas (Shapiro-Wilk): p-value = 0.85 (Normal)

Multikolinearitas (VIF): Semua VIF < 10 (Tidak ada multikolinearitas)

Heteroskedastisitas (Breusch-Pagan): p-value = 0.59 (Homoskedastik)

Autokorelasi (Durbin-Watson): 1.99 (Tidak ada autokorelasi)

C. Kesimpulan

- Model regresi linier berganda sangat baik ($R^2 = 98,3\%$) dalam menjelaskan variabilitas pendapatan harian.
- Jumlah pelanggan dan pengeluaran iklan berpengaruh positif signifikan terhadap pendapatan harian.
- Tingkat diskon berpengaruh negatif signifikan; peningkatan diskon justru menurunkan pendapatan harian.
- Seluruh uji asumsi klasik terpenuhi, sehingga model valid digunakan untuk pengambilan keputusan.
- Rekomendasi: Fokus pada peningkatan jumlah pelanggan, pengelolaan anggaran iklan, serta bijak dalam pemberian diskon.