


<p>Nama: Mohammad Afif R. Lingkeh NIM: 06500240000</p>	 UNIVERSITAS TRISAKTI	<p>MODUL 9 STATISTIKA</p> <p>Nama Dosen: Joko Riyono</p>
<p>Hari/Tanggal: Jumat, 23 Mei 2025</p>	<p>PRAKTIKUM STATISTIKA</p>	<p>Nama Aslab:</p> <ol style="list-style-type: none"> 1. Tarum Widyasti (064002200027) 2. Kharisma Maulida (064002200024)

Modul 9

Regresi Linier Sederhana dan Berganda

Teori Singkat

Pada regresi linier akan dibicarakan masalah pendugaan atau peramalan sebuah variabel dependen Y dengan sebuah variabel independen X yang telah diketahui nilainya. Model persamaan linier yang digunakan di sini adalah :

$$\hat{y} = a + bx$$

Regresi linier berganda Jika variabel dependen-nya dihubungkan dengan lebih dari satu variabel independen, maka persamaan yang dihasilkan adalah persamaan regresi linier berganda (*multiple linier regression*). Dalam hal ini kita membatasi pada kasus dua peubah bebas X_1 dan X_2 saja. Dengan hanya dua peubah bebas, persamaan regresi contohnya menjadi :

$$\hat{y} = b_0 + b_1x_1 + b_2x_2$$

Salah satu ukuran kebaikan model adalah dengan melihat koefisien determinasi R^2 yang menyatakan proporsi keragaman variabel Y yang dapat dijelaskan oleh variabel X. Namun penggunaan yang lebih baik adalah dengan menggunakan nilai **R-Sq(adj)**, yang merupakan nilai estimasi yang tidak bias (*unbiased estimate*) dari populasi.

ELEMEN KOMPETENSI I

No	X	Y
1	40	385
2	20	400
3	25	395
4	20	365
5	30	475
6	50	440

Misalkan ingin dilakukan pendugaan terhadap nilai penjualan dalam USD (variabel Y) berdasarkan nilai biaya iklan yang dikeluarkan dalam USD (variabel X) di suatu perusahaan. Data sampel dalam 12 bulan terakhir adalah sebagai berikut :

- buatlah persamaan regresi untuk menduga penjualan mingguan (Y) berdasarkan pengeluaran iklan (X).
- Hitunglah R-square (Koefisien determinasi) dan korelasinya.

R Studio:

```
df_nama=read.delim("clipboard")
View(df_nama)
model_reg=lm(df_nama$Y~df_nama$X)
summary(model_reg)
```

Console Terminal x Background Jobs x

R 4.4.3 · ~/

```
> df_afif=read.delim("clipboard")
> View(df_afif)
> model_reg=lm(df_afif$Y~df_afif$X)
> summary(model_reg)
```

Call:
lm(formula = df_afif\$Y ~ df_afif\$X)

Residuals:

1	2	3	4	5	6
-37.399	4.653	-7.110	-30.347	66.127	4.075

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	368.295	50.184	7.339	0.00184 **
df_afif\$X	1.353	1.534	0.882	0.42760

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

Residual standard error: 41.17 on 4 degrees of freedom
Multiple R-squared: 0.1628, Adjusted R-squared: -0.04648
F-statistic: 0.7779 on 1 and 4 DF, p-value: 0.4276

> |

Nilai R Squarenya adalah 0.1628
Korelasi nya adalah 0.40350105

Deskripsi (minimal 4 baris)

Berdasarkan output regresi linear sederhana dari RStudio, diperoleh persamaan $Y=368.295+1.353X$, yang menunjukkan bahwa setiap penambahan 1 USD pada biaya iklan diperkirakan meningkatkan penjualan sebesar 1.353 USD. Namun, koefisien regresi tersebut tidak signifikan secara statistik ($p\text{-value} = 0.4276 > 0.05$), yang berarti tidak terdapat cukup bukti untuk menyimpulkan adanya hubungan yang signifikan antara biaya iklan dan penjualan pada tingkat kepercayaan 95%. Nilai R-squared yang rendah (0.1628) juga menunjukkan bahwa hanya sekitar 16% variasi penjualan dapat dijelaskan oleh biaya iklan, sementara sisanya dipengaruhi oleh faktor lain. Dengan demikian, meskipun terdapat kecenderungan peningkatan penjualan seiring peningkatan biaya iklan, hubungan tersebut tidak signifikan secara statistik berdasarkan data yang tersedia.

Python:

```
import numpy as np
import pandas as pd

# Data
data = {
    "X": [40, 20, 25, 20, 30, 50],
    "Y": [385, 400, 395, 365, 475, 440]
}
df = pd.DataFrame(data)

# Menghitung rata-rata
mean_X = np.mean(df["X"])
mean_Y = np.mean(df["Y"])

# Menghitung b1 (slope)
df['XY'] = df['X'] * df['Y']
df['X^2'] = df['X'] ** 2
b1 = (sum(df['XY']) - len(df) * mean_X * mean_Y) / (sum(df['X^2']) - len(df) * mean_X ** 2)

# Menghitung b0 (intersep)
b0 = mean_Y - b1 * mean_X

# Persamaan regresi
regression_eq = f"Y = {b0:.2f} + {b1:.2f}X"
```

```
# Menghitung koefisien korelasi (r)
correlation_matrix = np.corrcoef(df["X"], df["Y"])
r = correlation_matrix[0, 1]

# Menghitung koefisien determinasi (R^2)
R2 = r ** 2

# Menampilkan hasil dengan deskripsi
print(f"Rata-rata X: {mean_X:.2f}")
print(f"Rata-rata Y: {mean_Y:.2f}")
print(f"Nilai intersep (b0): {b0:.2f}")
print(f"Nilai kemiringan (b1): {b1:.2f}")
print(f"Koefisien korelasi (r): {r:.2f}")
print(f"Koefisien determinasi (R^2): {R2:.2f}")
print(f"Persamaan regresi: {regression_eq}")

# Menampilkan DataFrame
print("\nDataFrame:")
print(df)
```

Output:

```
➡ Rata-rata X: 30.83
Rata-rata Y: 410.00
Nilai intersep (b0): 368.29
Nilai kemiringan (b1): 1.35
Koefisien korelasi (r): 0.40
Koefisien determinasi (R^2): 0.16
Persamaan regresi: Y = 368.29 + 1.35X
```

DataFrame:

	X	Y	XY	X^2
0	40	385	15400	1600
1	20	400	8000	400
2	25	395	9875	625
3	20	365	7300	400
4	30	475	14250	900
5	50	440	22000	2500

Deskripsi (minimal 4 baris)

Berdasarkan hasil analisis regresi linear sederhana menggunakan Python terhadap data biaya iklan dan penjualan, diperoleh persamaan regresi $Y = 368.29 + 1.35X$, yang menunjukkan bahwa setiap kenaikan 1 USD pada biaya iklan diperkirakan akan meningkatkan penjualan sebesar 1.35 USD. Rata-rata biaya iklan adalah sebesar 30.83 USD dan rata-rata penjualan sebesar 410 USD. Koefisien korelasi sebesar 0.40 menunjukkan hubungan positif yang lemah antara biaya iklan dan penjualan, dan nilai koefisien determinasi (R^2) sebesar 0.16 mengindikasikan bahwa hanya 16% variasi dalam penjualan yang dapat dijelaskan oleh variasi dalam biaya iklan. Dengan demikian, meskipun terdapat kecenderungan hubungan positif antara biaya iklan dan penjualan, hubungan ini tergolong lemah dan faktor-faktor lain kemungkinan besar turut memengaruhi besarnya penjualan.

Excel:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		x	y	x^2	xy			(x-xbar)	(y-ybar)	(x-xbar)(y-ybar)	(x-xbar)^2		y_duga	(y_duga - ybar)^2	(y-ybar)^2	
2		40	385	1600	15400			9.17	-25	-229.17	84.03		422.3988439	153.7313308	625	
3		20	400	400	8000			-10.83	-10	108.33	117.36		395.3468208	214.7156604	100	
4		25	395	625	9875			-5.83	-15	87.50	34.03		402.1098266	62.25483645	225	
5		20	365	400	7300			-10.83	-45	487.50	117.36		395.3468208	214.7156604	2025	
6		30	475	900	14250			-0.83	65	-54.17	0.69		408.8728324	1.270506866	4225	
7		50	440	2500	22000			19.17	30	575.00	367.36		435.9248555	672.0981322	900	
8	Total	185	2460	6425	76825					975.00	720.83			1318.786127	8100	
9														SSR	SST	
10	x_bar	30.83														
11	y_bar	410														
12	n	6														
13	b1	1.352601						b1		1.352601156			rsquare	0.162813102		
14	b0	368.2948						b0		368.2947977			correlation	0.403501056		
15																
16	y_duga =	358.2948+1.3526x								NAMA	Mohammad Afif R. Lingkeh					
17	artinya penambahan satu satuan x akan meningkatkan 1,3526 satuan y									NIM	65002400002					
18																
19																

Nilai R Squarenya adalah 0.1628
Korelasi nya adalah 0.40350105

Deskripsi (minimal 4 baris)

Berdasarkan hasil perhitungan menggunakan Excel terhadap data biaya iklan (X) dan penjualan (Y), diperoleh bahwa rata-rata biaya iklan adalah sekitar 30,83 dan rata-rata penjualan sebesar 410. Persamaan regresi yang dihasilkan adalah $Y = 368,29 + 1,35X$, yang berarti setiap kenaikan 1 USD pada biaya iklan diperkirakan akan meningkatkan penjualan sebesar 1,35 USD. Analisis lebih lanjut menunjukkan bahwa hanya sebagian kecil variasi penjualan yang dapat dijelaskan oleh variasi biaya iklan, sebagaimana terlihat dari perhitungan selisih kuadrat total dan prediksi. Ini menunjukkan bahwa meskipun terdapat hubungan positif antara biaya iklan dan penjualan, pengaruh biaya iklan terhadap penjualan tidak dominan, dan kemungkinan terdapat faktor lain yang lebih berpengaruh terhadap tingkat penjualan.

TUGAS

Delapan orang lulusan dipilih secara random dan ditanyakan berapa nilai IPK kelulusan (X) serta total gaji pertama kali (Y) dalam jutaan rupiah. Data yang diperoleh sebagai berikut:

Nama	X	Y
Amir	2,8	5,4
Agus	2,5	5,1
Charlie	3,5	7,2
Debi	3,1	6,2
Faishal	3,0	6,0
Jojo	3,8	7,5
Kamal	3,3	6,8
Caca	3,5	8,9

Pertanyaan:

- Hitung nilai intersep (b0) dan slope/koeffisien regresi (b1). Jelaskan perhitungan manualnya
- Tuliskan persamaan regresi linear sederhana dan interpretasikan nilai b1 dalam regresi ini.
- Hitung nilai koefisien determinasi dan koefisien korelasi serta interpretasikan nilainya. Jelaskan perhitungan manualnya

R Studio:

```
> df_afif=read.delim("clipboard")
> View(df_afif)
> model_reg=lm(df_afif$Y~df_afif$X)
> summary(model_reg)

Call:
lm(formula = df_afif$Y ~ df_afif$X)

Residuals:
    Min       1Q   Median       3Q      Max
-0.66967 -0.23146 -0.19354 -0.04362  1.48078

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.3360     1.9930  -0.670   0.52756
df_afif$X      2.5015     0.6205   4.031   0.00687 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6934 on 6 degrees of freedom
Multiple R-squared:  0.7304,    Adjusted R-squared:  0.6854
F-statistic: 16.25 on 1 and 6 DF,  p-value: 0.006869

> |
```

Deskripsi (minimal 4 baris)

Hasil regresi linear menunjukkan bahwa IPK berpengaruh signifikan terhadap gaji pertama lulusan, dengan persamaan $Y = -1,336 + 2,5015X$. Koefisien IPK (2,5015) bermakna bahwa setiap kenaikan 1 poin IPK diperkirakan meningkatkan gaji sebesar Rp 2,5015 juta. Nilai R^2 sebesar 0,7304

menunjukkan bahwa 73% variasi gaji dapat dijelaskan oleh IPK, sementara sisanya dipengaruhi faktor lain. Uji signifikansi menghasilkan p-value 0,00687 ($< 0,01$), yang berarti model ini signifikan secara statistik dan hubungan antara IPK dan gaji bersifat kuat dan positif.

Python:

```

⇒ Rata-rata X: 3.19
   Rata-rata Y: 6.64
   Nilai intersep (b0): -1.34
   Nilai kemiringan (b1): 2.50
   Koefisien korelasi (r): 0.85
   Koefisien determinasi (R^2): 0.73
   Persamaan regresi: Y = -1.34 + 2.50X
    
```

DataFrame:

	X	Y	XY	X^2
0	2.8	5.4	15.12	7.84
1	2.5	5.1	12.75	6.25
2	3.5	7.2	25.20	12.25
3	3.1	6.2	19.22	9.61
4	3.0	6.0	18.00	9.00
5	3.8	7.5	28.50	14.44
6	3.3	6.8	22.44	10.89
7	3.5	8.9	31.15	12.25

Deskripsi (minimal 4 baris)

Hasil analisis regresi menggunakan Python menunjukkan bahwa terdapat hubungan positif yang kuat antara IPK (X) dan gaji pertama (Y) lulusan, dengan rata-rata IPK sebesar 3,19 dan rata-rata gaji sebesar Rp 6,64 juta. Persamaan regresi yang dihasilkan adalah $Y = -1,34 + 2,50X$, yang berarti setiap kenaikan IPK 1 poin diperkirakan meningkatkan gaji pertama sebesar Rp 2,5 juta. Koefisien korelasi sebesar 0,85 menunjukkan hubungan yang sangat kuat antara IPK dan gaji, sementara koefisien determinasi (R^2) sebesar 0,73 menunjukkan bahwa 73% variasi gaji dapat dijelaskan oleh IPK, sisanya dipengaruhi oleh faktor lain di luar model. Data yang digunakan mencakup delapan observasi, dan perhitungan XY serta X^2 mendukung hasil tersebut.

Excel:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Nama	X	Y	x^2	xy			(x-xbar)	(y-ybar)	(x-xbar)(y-ybar)	(x-xbar)^2		y_duga	(y_duga-ybar)^2	(y-ybar)^2	
2	Amir	2.8	5.4	7.84	15.12			-0.39	-1.2375	0.48	0.15		5.668168168	0.9396042	1.53140625	
3	Agus	2.5	5.1	6.25	12.75			-0.69	-1.5375	1.06	0.47		4.917717718	2.957651098	2.36390625	
4	Charlie	3.5	7.2	12.25	25.2			0.31	0.5625	0.18	0.10		7.419219219	0.611084938	0.31640625	
5	Debi	3.1	6.2	9.61	19.22			-0.09	-0.4375	0.04	0.01		6.418618619	0.047909059	0.19140625	
6	Faishal	3	6	9	18			-0.19	-0.6375	0.12	0.04		6.168468468	0.219990578	0.40640625	
7	Jojo	3.8	7.5	14.44	28.5			0.61	0.8625	0.53	0.38		8.16966967	2.347543897	0.74390625	
8	Kamal	3.3	6.8	10.89	22.44			0.11	0.1625	0.02	0.01		6.918918919	0.079196608	0.02640625	
9	Caca	3.5	8.9	12.25	31.15			0.31	2.2625	0.71	0.10		7.419219219	0.611084938	5.11890625	
10	Total	25.5	53.1	82.53	172.38					3.12	1.25			7.814065315	10.69875	
11																
12																
13	x_bar	3.19						b1	2.501501502				rsquare	0.730371802		
14	y_bar	6.6375						b0	-1.336036036				correlation	0.854617927		
15	n	8														
16	b1	2.501502														
17	b0	-1.33604														
18										NAMA	Mohammad Afif R. Lingkeh					
19										NIM	65002400002					
20	y_duga =	-1,33604+2,501502x														
21	artinya penambahan satu satu satuan x akan meningkatkan 2,501502 satuan y															
22																

Deskripsi (minimal 4 baris)

Hasil analisis regresi menunjukkan bahwa terdapat hubungan positif dan kuat antara IPK dan gaji pertama lulusan, dengan persamaan regresi $Y = -1,336 + 2,502X$. Artinya, setiap kenaikan IPK 1 poin diperkirakan meningkatkan gaji pertama sebesar Rp 2,502 juta. Koefisien korelasi sebesar 0,85 menunjukkan hubungan yang kuat, dan nilai koefisien determinasi sebesar 73% mengindikasikan bahwa sebagian besar variasi gaji pertama dapat dijelaskan oleh IPK, sementara sisanya dipengaruhi faktor lain di luar model.

CEK LIST (✓)

1. Melakukan regresi linier sederhana dan berganda. (✓)

KESIMPULAN

Dari praktikum ini, saya mempelajari cara menghitung dan menginterpretasikan regresi linear sederhana dan berganda, termasuk mencari nilai intersep, koefisien regresi, koefisien korelasi, dan determinasi. Saya juga belajar menggunakan Excel, RStudio, dan Python untuk melakukan analisis regresi secara praktis. Selain itu, saya memahami pentingnya uji signifikansi dalam menentukan apakah hubungan antar variabel benar-benar berarti secara statistik.

FORM UMPAN BALIK

Elemen Kompetensi	Tingkat Kesulitan	Tingkat Ketertarikan	Waktu Penyelesaian (menit)
Melakukan regresi linier sederhana dan berganda.	1	4	120

Keterangan Tingkat Kesulitan

- 1: Sangat Mudah
- 2: Mudah
- 3: Biasa
- 4: Sulit
- 5: Sangat Sulit

Keterangan Tingkat Ketertarikan

- 1: Tidak Tertarik
- 2: Cukup Tertarik
- 3: Tertarik
- 4: Sangat Tertarik