

1. Jelaskan perbedaan antara **korelasi** dan **regresi linier**.

- Apa yang diukur oleh korelasi?
- Apa yang dihasilkan oleh analisis regresi?

2. Pada persamaan regresi linier sederhana

$$Y = a + bX$$

jelaskan makna:

- parameter **a (intersep)**
- parameter **b (koefisien regresi / slope)**

Gunakan contoh konteks: *pengalaman kerja (X)* dan *omzet penjualan (Y)*.

3. Apa yang dimaksud dengan **koefisien determinasi (R^2)**? Bagaimana cara menafsirkannya terhadap “kekuatan” pengaruh variabel X terhadap Y?
4. Jelaskan pengertian **selisih taksir standar (standard error of estimate)** dalam regresi linier. Mengapa nilai ini penting untuk menilai ketepatan model?
5. Uraikan secara runtut **langkah-langkah analisis regresi linier sederhana** mulai dari data mentah sampai interpretasi hasil (minimal 5 langkah utama).

B. Soal Perhitungan Manual

1.

X	1	2	3	4	5	6
Y	6	4	3	5	4	2

- A. Hitung persamaan regresi linier sederhana $Y = a + bX$
- B. Hitung nilai dugaan $\hat{Y}$ jika $X=8$
- C. Buat tabel yang memuat $Y, \hat{Y}, Y - \hat{Y}$, dan $(Y - \hat{Y})^2$
- D. Gunakan hasil butir (3) untuk menghitung **selisih taksir standar** (standard error of estimate) sesuai rumus pada slide.
- E. Interpretasikan makna nilai selisih taksir standar yang Anda dapatkan (apakah pencaran data terhadap garis regresi cukup besar atau kecil?).

2.

No	X (tahun)	Y (ribu rupiah)
1	2	5
2	3	8
3	2	8
4	5	7
5	6	11
6	1	3
7	4	10
8	1	4

Diminta:

1. Buat tabel bantu yang memuat X , Y , X^2 , Y^2 , XY beserta jumlah (Σ).
2. Hitung nilai **a** dan **b** secara manual.
3. Tulis persamaan regresi linier **Y atas X**.
4. Hitung nilai dugaan omzet ($\hat{Y}$) jika pengalaman kerja karyawan adalah **3,5 tahun**.
5. Jelaskan secara singkat arti tanda dan besar nilai b yang Anda peroleh.

c. Tugas Pratikum

C1. Menjalankan Regresi Linier di Google Colab

1. Buat notebook baru di **Google Colab**.
2. Tuliskan dan jalankan kode berikut, lalu sesuaikan dengan data Anda sendiri:

```
# 1. Import pustaka yang dibutuhkan
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_squared_error

# 2. Input data (X = pengalaman kerja, Y = omzet)
X = np.array([2, 3, 2, 5, 6, 1, 4, 1]).reshape(-1, 1) # bentuk kolom
Y = np.array([5, 8, 8, 8, 7, 11, 3, 10, 4])

# 3. Membuat dan melatih model
model = LinearRegression()
model.fit(X, Y)

# 4. Menampilkan parameter regresi
a = model.intercept_ # konstanta
b = model.coef_[0]   # koefisien regresi
print("Intercept (a) =", a)
print("Slope (b)    =", b)

# 5. Menghitung prediksi dan evaluasi
Y_pred = model.predict(X)

# Koefisien determinasi (R^2)
r2 = r2_score(Y, Y_pred)

# SSE (sum of squared errors) dan MSE
sse = np.sum((Y - Y_pred) ** 2)
mse = mean_squared_error(Y, Y_pred)

print("R^2 =", r2)
print("SSE =", sse)
print("MSE =", mse)

# 6. Prediksi untuk pengalaman kerja 3.5 tahun
x_new = np.array([[3.5]])
y_new = model.predict(x_new)
print("Perkiraan omzet jika pengalaman 3.5 tahun =", y_new[0])
```

Jawab pertanyaan berikut berdasarkan output program:

1. Bandingkan nilai **a** dan **b** yang dihasilkan oleh program dengan hasil **perhitungan manual** Anda di Soal 1. Apakah sama (perbedaan hanya di pembulatan)? Jelaskan.
2. Berapa nilai R^2 dari model? Apa artinya terhadap pengaruh pengalaman kerja terhadap omzet penjualan?
3. Bandingkan nilai **SSE** (jumlah kuadrat galat) dari output program dengan hasil perhitungan manual pada Soal 1 (jika dihitung).
4. Ubah sendiri data X dan Y (misalnya menambah 2–3 data baru), jalankan ulang kode, lalu jelaskan perubahan yang terjadi pada nilai b dan R^2 .

D. Tugas Mencari Jurnal (Literature Review Sederhana)

1. Gunakan **Google Scholar** atau portal jurnal lainnya.
2. Carilah **minimal 2 jurnal** yang menggunakan **regresi linier sederhana** atau **regresi linier berganda**, dengan ketentuan:
 - 1 jurnal nasional terakreditasi (misal SINTA).
 - 1 jurnal internasional (Scopus/WoS atau jurnal bereputasi lainnya).
3. Untuk setiap jurnal, tuliskan ringkasan (± 1 paragraf) yang memuat:
 - Judul dan identitas jurnal (nama jurnal, tahun, penulis utama).
 - Tujuan penelitian.
 - Variabel bebas (X) dan variabel terikat (Y).
 - Model regresi yang digunakan (tuliskan persamaan jika tersedia).
 - Nilai R^2 atau indikator lain yang dilaporkan, beserta interpretasinya.
 - Bidang/topik penelitian (misal: pemasaran, keuangan, kesehatan, pendidikan, dsb).
4. Tulis sitasi dalam format sederhana, misalnya:
 - *Nama Belakang, Inisial. (Tahun). Judul Artikel. Nama Jurnal, Volume(Nomor), halaman.*
5. Lampirkan **tautan (URL)** menuju jurnal tersebut di laporan / lembar jawaban Anda.