

# Jawaban Tugas 2 STDA4101-24.1 (1 dan 2)

November 19, 2024

---

Nama: Moh Evan Alsadik | Nim: 053779188 | STDA4101

```
[3]: from scipy.stats import linregress
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
file_tugas = "Tugas 2 STDA4101-2024.1.xlsx"
tugas_df = pd.read_excel(file_tugas)
# tugas_df.describe()
# tugas_df.info()
usia = tugas_df["AGE"]
perkembangan_penyakit = tugas_df["Y"]
print("\n")
```

Berdasarkan data ukuran kuantitatif perkembangan penyakit setelah satu tahun amatan (Y) dan usia dalam tahun (AGE) terlampir, lakukan analisis statistik deskriptif dengan cara:

---

## 0.0.1 Buatlah visualisasi data bivariat menggunakan Scatter Plot

```
[8]: # Scatter Plot
def scatter_plot():
    plt.figure(figsize=(12, 6))
    plt.scatter(
        usia,
        perkembangan_penyakit,
        color='dodgerblue',
        alpha=0.7,
        edgecolor='black',
        s=100,
        label="Data"
    )
```

```

# Label dan Grid
def label_grid():
    # Label dan judul
    plt.title("Grafik Scatter Plot untuk Usia vs Perkembangan Penyakit",
              fontsize=16,
              weight='bold')
    plt.xlabel("Usia (dalam Tahun)", fontsize=14)
    plt.ylabel("Perkembangan Penyakit", fontsize=14)

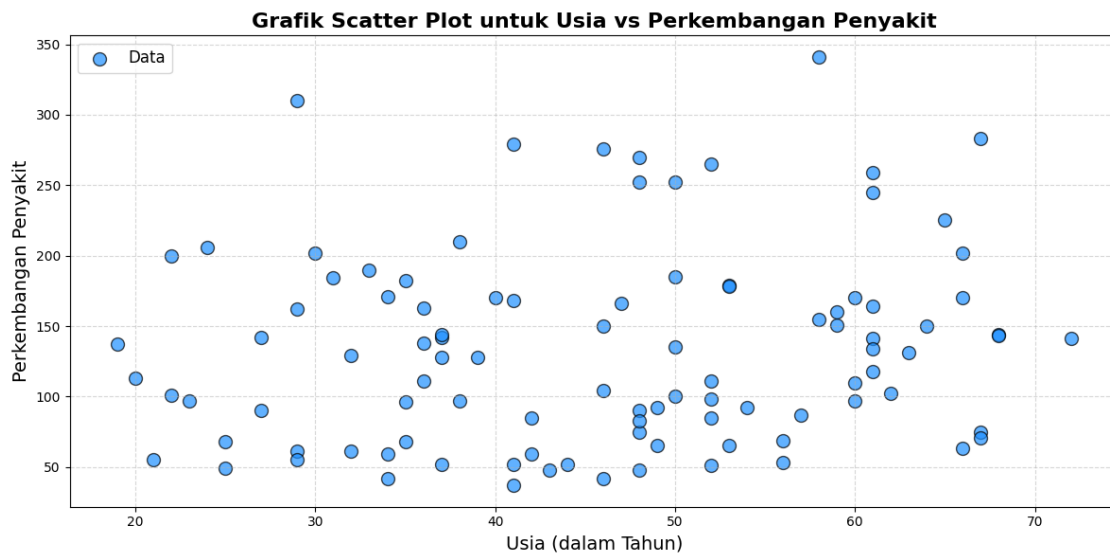
    # Grid dan legenda
    plt.grid(True, linestyle='--', alpha=0.5)
    plt.legend(loc="upper left", fontsize=12)

scatter_plot()
label_grid()

# Tampilkan plot
plt.tight_layout()
plt.show()

print("\n")
print("\n")
print("\n")

```



---

## 0.0.2 Tentukan model persamaan regresi linear

```
[7]: # Hitung regresi linear
slope, intercept, r_value, p_value, std_err = linregress(usia,
    ↳ perkembangan_penyakit)

# Grafik Scatter Plot dengan Garis Regresi
scatter_plot()

# Susun data untuk garis regresi
sorted_usia = sorted(usia)
regresi_y = slope * pd.Series(sorted_usia) + intercept # Hitung nilai Y untuk
    ↳ setiap nilai X

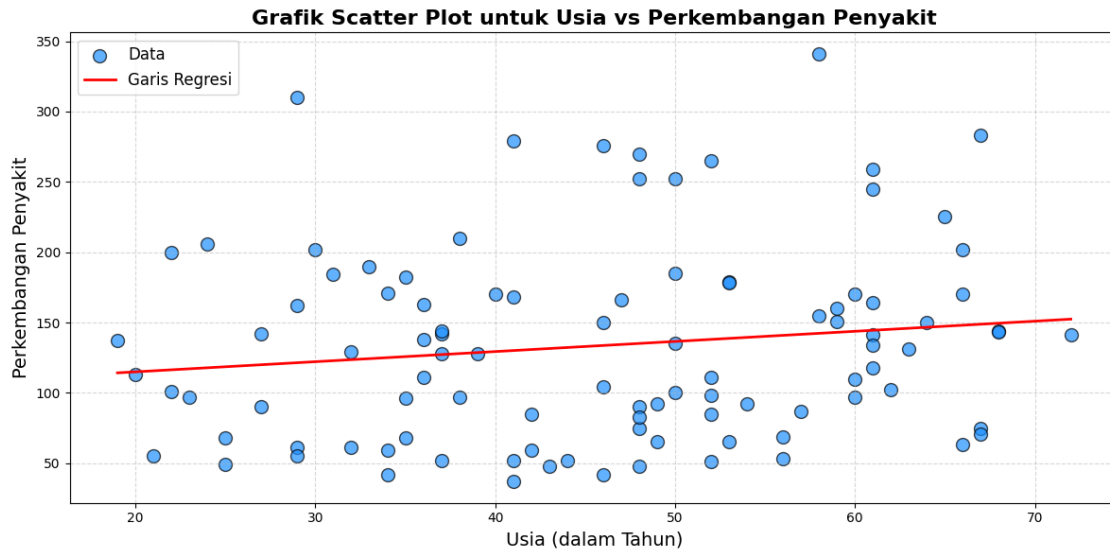
# Garis regresi
plt.plot(sorted_usia, regresi_y, color="red", linewidth=2, label="Garis
    ↳ Regresi")

label_grid()

# Tampilkan plot
plt.tight_layout()
plt.show()
print("\n")

# Menampilkan hasil regresi
print("Persamaan Regresi Linear:")
print(f"Y = {intercept:.2f} + {slope:.2f} * X")
print(f"Koefisien Korelasi (r): {r_value:.2f}") # menunjukkan arah hubungan
    ↳ linier antara dua variabel (-/+)
print(f"p-value: \t\t{p_value:.4f}") # menguji signifikansi hubungan antar
    ↳ variabel
print(f"Standard Error: \t{std_err:.2f}") # mengukur tingkat ketidakpastian
    ↳ dalam estimasi (slope)

print("\n")
print("\n")
```



Persamaan Regresi Linear:

$$Y = 100.56 + 0.72 * X$$

Koefisien Korelasi (r): 0.14

p-value: 0.1572

Standard Error: 0.51