

Modul Lanjutan Analisis Data Menggunakan Python

A. Identitas Modul

Topik Modul	Analisis Data Penjualan Menggunakan Python
Bidang Keahlian	Rekayasa Perangkat Lunak / Data Science
Topik	Data Cleaning, Analisis Deskriptif, dan Visualisasi Data
Perangkat Lunak	Python (Anaconda / Google Colab / Jupyter Notebook)
Durasi Kegiatan	± 3 jam praktik

B. Tujuan Pembelajaran

- Menjelaskan fungsi utama library pandas, numpy, matplotlib, dan seaborn.
- Melakukan proses import, pembersihan, dan analisis data menggunakan Python.
- Membuat visualisasi data berupa grafik batang dan grafik garis.
- Menginterpretasikan hasil analisis sederhana untuk pengambilan keputusan.

C. Konsep Dasar Teori

Analisis data adalah proses mengumpulkan, membersihkan, dan menginterpretasikan data untuk memperoleh informasi yang berguna bagi pengambilan keputusan. Berikut beberapa library utama Python untuk analisis data:

Library	Kegunaan Utama
NumPy	Operasi matematika dan array
Pandas	Mengelola data dan membaca data berbasis tabel (DataFrame)
Matplotlib	Visualisasi data dalam bentuk grafik
Seaborn	Visualisasi statistik yang menarik dan informatif

D. Langkah-Langkah Praktikum

Langkah 1 – Import Library

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Langkah 2 – Import Data dari File CSV

```
df =
pd.read_csv("Data_Penjualan_Toko_Online.csv")
print(df.head())
```

Langkah 3 – Mengecek dan Membersihkan Data

```
print(df.info())
print(df.isnull().sum())
# Jika ada nilai kosong
df = df.dropna()
```

Langkah 4 – Analisis Deskriptif

```
print("Statistik Deskriptif:")
print(df.describe())

print("\nProduk Terlaris:")
print(df.groupby("Produk") ["Jumlah"].sum().sort_values(ascending= False))
```

Langkah 5 – Visualisasi Data

```
plt.figure(figsize=(8,5))
sns.barplot(data=df, x="Produk", y="Pendapatan",
estimator=sum, ci=None)
plt.title("Total Pendapatan per Produk")
plt.show()
```

```
plt.figure(figsize=(10,5))
sns.lineplot(data=df, x="Tanggal", y="Pendapatan",
marker="o") plt.title("Tren Pendapatan Mingguan")
plt.show()
```

Langkah 6 – Menyimpan Hasil Analisis

```
df.to_csv("hasil_analisis_penjualan.csv",
index=False) print("Data hasil analisis telah
disimpan sebagai hasil_analisis_penjualan.csv")
```

E. Praktik Mandiri

- Tambahkan kolom baru bernama 'Diskon' sebesar 10% dari harga untuk semua produk.
- Hitung kolom baru 'Pendapatan Bersih' = Pendapatan - (Pendapatan × Diskon).
- Buat grafik batang untuk menampilkan pendapatan bersih tiap produk.
- Simpan hasilnya dalam file analisis_diskon.csv.

F. Evaluasi Pemahaman

No	Pertanyaan	Jawaban Singkat
1	Apa fungsi library pandas dalam analisis data?	untuk membaca dan mengolah file csv
2	Apa perbedaan antara matplotlib dan seaborn?	matplotlib library untuk mengvisualisasi data, seaborn library untuk mempercantik visual data
3	Mengapa perlu melakukan data cleaning sebelum analisis?	untuk membersihkan data data yang bermasalah
4	Fungsi apa yang digunakan untuk menampilkan 5 data pertama dari DataFrame?	df.head()
5	Bagaimana cara menyimpan data hasil analisis ke file CSV?	dengan menggunakan method "to_csv"

G. Kesimpulan

Dalam modul ini, Anda telah belajar langkah-langkah dasar analisis data menggunakan Python dengan data CSV, mulai dari mengimpor file, membersihkan, menganalisis, hingga memvisualisasikan hasilnya. Keterampilan ini merupakan fondasi penting untuk menuju bidang data analytics dan *machine learning*.