INFORMATIKA · KELAS X · FASE E

LEMBAR KERJA PESERTA DIDIK

# LKPD

## BAB 6
# Analisis Data

Kumpulkan, rapikan, dan ceritakan datamu dengan Python.





```
import pandas as pd
df = scrape(web)
df.plot.bar()
# insight!
>>> Halo, data!
```

| Nama | Kelas / No. Absen |
|---|---|
| | |
| **Kelompok** | **Sekolah** |
| | |

LIVEWORKSHEETS

# LEMBAR KERJA PESERTA DIDIK (LKPD)

*Mata Pelajaran: Informatika • Kelas X SMA • Fase E • Kurikulum Merdeka*

**BAB 6 | ANALISIS DATA**

| | |
|---|---|
| **Nama Lengkap :** ..... | **Kelas / Fase :** ..... |
| **No. Presensi / Kelompok :** ..... | **Hari, Tanggal :** ..... |

## A Tujuan Pembelajaran

Setelah mengerjakan LKPD ini, kalian diharapkan bisa:

1. menjelaskan alur analisis data dan perkakas yang dipakai (Google Colab, Python, dan library);
2. menulis dan menjalankan program Python sederhana di Google Colab;
3. mengoleksi data dari halaman web dengan web scraping dan menampilkannya sebagai tabel;
4. menilai apakah suatu pengambilan data pantas dilakukan dari sisi privasi dan keamanan;
5. merapikan data lalu memilih dan membuat diagram yang tepat.

## B Petunjuk Belajar

1. Siapkan laptop atau komputer, koneksi internet, dan akun Google.
2. Tulis identitasmu pada kolom yang tersedia di atas.
3. Bentuk kelompok 3–4 orang. Bagi tugas: satu orang mengetik kode, satu orang mencatat jawaban, yang lain membaca soal dan mengecek hasil. **Gantian** setiap sub bab.
4. Baca kotak **Ringkasan Materi** dulu sebelum mengerjakan soal di bawahnya.
5. Perhatikan label berwarna di depan setiap soal, karena cara mengerjakannya berbeda-beda.
6. Ketik sendiri semua kode di Colab, jangan hanya salin-tempel. Kalau muncul error, baca pesannya, catat, lalu coba perbaiki bersama dulu sebelum bertanya ke guru.
7. Tulis jawaban dengan kalimat sendiri, lalu periksa kembali sebelum dikumpulkan.

## C Peta Belajar

Bab ini punya tiga sub bab. Pantau perjalananmu di sini.

LIVEWORKSHEETS

## D Kenali Jenis Soalnya

Setiap soal punya label warna. Inilah artinya:

| | |
|---|---|
| CEK CEPAT | Jawab singkat: pilihan ganda atau benar-salah |
| TEBAK | Tebak hasilnya dulu, baru jalankan untuk membuktikan |
| PRAKTIK | Ketik dan jalankan kode di Google Colab |
| BUG | Cari dan perbaiki kesalahan pada kode |
| BACA KODE | Jelaskan apa yang dilakukan setiap baris kode |
| JODOHKAN | Pasangkan dua hal yang berhubungan |
| DISKUSI | Bahas bersama kelompok, lalu catat kesepakatan |
| MISI | Tugas mini yang menghasilkan sesuatu |

1

SUB BAB 1

# Pengenalan Perkakas Analisis Data

*Kenali alur analisis data, Google Colab, dan dasar Python.*

### RINGKASAN MATERI

**Analisis data** adalah proses mengolah data mentah menjadi informasi yang berguna untuk mengambil keputusan. Alurnya: mengambil data, memproses (merapikan) data, menyajikan data dalam tabel atau diagram, menganalisis pola, menginterpretasi hasil, lalu menyimpulkan.

Kalau datanya banyak atau harus diambil berulang-ulang, pekerjaan ini jauh lebih cepat dengan bantuan program komputer. Perkakas yang kita pakai di bab ini:

| Perkakas | Fungsi |
|---|---|
| **Google Colab** | Tempat menulis dan menjalankan program Python lewat browser, gratis dan tanpa instalasi |
| **Python** | Bahasa pemrograman yang dipakai untuk mengolah data |
| **requests dan BeautifulSoup** | Mengambil dan membaca kode HTML halaman web |
| **pandas** | Menyusun data menjadi tabel (DataFrame) |
| **plotly** | Membuat diagram dari data |

LIVEWORKSHEETS

## Bagian 1 · Alur dan perkakas

**Soal 1.1** CEK CEPAT **Benar atau salah? Beri tanda ✔ pada kolom yang sesuai.**

| No | Pernyataan | Benar | Salah |
|---|---|---|---|
| 1 | Analisis data hanya bisa dilakukan dengan komputer. | ☐ | ☐ |
| 2 | Data mentah sering perlu dirapikan sebelum divisualisasikan. | ☐ | ☐ |
| 3 | Menyimpulkan dilakukan sebelum data diambil. | ☐ | ☐ |
| 4 | Diagram membantu kita melihat pola yang sulit terlihat di tabel angka. | ☐ | ☐ |
| 5 | Program komputer bisa mengambil data dari web secara otomatis. | ☐ | ☐ |

**Soal 1.2** CEK CEPAT **Urutkan alurnya. Enam tahap ini sengaja diacak, beri nomor 1–6.**

| Tahap | Urutan ke- |
|---|---|
| Menginterpretasi hasil | |
| Mengambil data | |
| Menyimpulkan dan mengambil keputusan | |
| Menyajikan data dalam tabel atau diagram | |
| Memproses atau merapikan data | |
| Menganalisis pola data | |

LIVEWORKSHEETS

### Soal 1.3 JODOHKAN Pasangkan perkakas dengan fungsinya. Tulis huruf yang sesuai.

| Perkakas | Huruf | Pilihan fungsi |
|---|---|---|
| 1. Google Colab | | a. Mengambil dan membaca kode HTML halaman web |
| 2. pandas | | b. Tempat menulis dan menjalankan Python lewat browser |
| 3. requests dan BeautifulSoup | | c. Membuat diagram dari data |
| 4. plotly | | d. Menyusun data dalam tabel |

## Bagian 2 · Mengenal Google Colab

### Soal 1.4 MISI Colab pertamamu

Buka https://colab.research.google.com/, pilih **File > New notebook**, ganti namanya menjadi `LKPD6_NamaKalian`, lalu jalankan kode ini dengan tombol **Run** (ikon segitiga di kiri sel).

```
print('Halo, nama saya ...')
```

Ganti titik-titik dengan nama kalian, lalu tulis keluarannya:

**Keluaran:**

Jelajahi tampilan Colab, lalu tuliskan fungsi tiga bagiannya (misalnya Menubar, Panel kiri, Cell toolbar):

| Bagian Colab | Fungsinya |
|---|---|
| | |
| | |
| | |

LIVEWORKSHEETS

## Bagian 3 · Dasar Python

> **RINGKASAN MATERI**
> Python di sini hanya alat. Enam konsep berikut sudah cukup untuk bab ini.

| Konsep | Fungsi | Contoh |
|---|---|---|
| **print** | Menampilkan data | `print('Halo')` |
| **Variabel** | Wadah untuk menyimpan nilai | `umur = 16` |
| **Array (list)** | Banyak nilai dalam satu variabel, indeks mulai dari 0 | `nilai = [10, 9, 8]` |
| **for** | Mengulang perintah untuk setiap elemen | `for n in nilai:` |
| **try-except** | Mencoba perintah, dan jika gagal lanjut ke cadangan | `try:` lalu `except:` |
| **import** | Memakai library (kumpulan kode siap pakai) | `import pandas as pd` |

### Soal 1.5 TEBAK Sebelum dijalankan, tebak dulu!

Tulis tebakan, jalankan, lalu isi hasil sebenarnya. Petunjuk: `len()` menghitung banyak elemen, dan tanda `*` pada teks berarti mengulang.

| Kode | Tebakanku | Hasil sebenarnya |
|---|---|---|
| `print(2 + 3)` | | |
| `print('2' + '3')` | | |
| `print('Ha' * 3)` | | |
| `nilai = [10, 9, 8]` lalu `print(nilai[2])` | | |
| `print(len([5, 6, 7, 8]))` | | |

**Berapa tebakan yang tepat? ..... dari 5. Apa yang kalian pelajari dari tebakan yang meleset?**

LIVEWORKSHEETS

## Soal 1.6 CEK CEPAT Pilihan ganda. Lingkari huruf jawaban yang paling tepat.

1. Jika `nilai = [7, 8, 9]`, maka `print(nilai[0])` menghasilkan …

   **A** 7 **B** 8
   **C** 9 **D** error

2. Library Python yang dipakai untuk membuat tabel data (DataFrame) adalah …

   **A** plotly **B** pandas
   **C** BeautifulSoup **D** requests

3. Fungsi utama try-except adalah …

   **A** mempercepat program **B** mengulang perintah
   **C** mencegah program berhenti saat terjadi error **D** menghapus variabel

## Soal 1.7 BUG Detektif kode: ada tiga kesalahan!

Program ini seharusnya menjumlahkan semua nilai, tetapi ada **tiga kesalahan**. Temukan, lalu tulis versi yang benar.

```
nilai = [80, 75, 90, 85]
total = 0
for n in nilai
print(n)
    total = total + n
print('Total:', Total)
```

| Kesalahan ke- | Di baris | Seharusnya |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |

## Soal 1.8 PRAKTIK Lengkapi kodenya

Isi bagian bergaris bawah supaya program mencetak rata-rata nilai, lalu jalankan untuk membuktikan.

```
nilai = [80, 75, 90, 85]
total = 0
for n in _________:
    total = total ____ n
rata = total / _________
print('Rata-rata:', rata)
```

LIVEWORKSHEETS

**Isian (urut dari atas) dan hasil yang tercetak:**

## Soal 1.9 MISI Tantangan bintang

Ubah data menjadi `nilai = [80, 'absen', 90, 85]`. Tulis versi program yang tetap berjalan dengan menambahkan try-except, lalu tuliskan keluarannya.

**Kode kalian:**

**Keluarannya:**

LIVEWORKSHEETS

SUB BAB 2

# 2 Koleksi Data

*Ambil data dari web secara otomatis, dan pahami batas etikanya.*

### RINGKASAN MATERI

Koleksi data adalah tahap mengumpulkan data yang akan dianalisis. Salah satu caranya adalah **web scraping**, yaitu mengambil data tertentu dari halaman web secara otomatis lalu menyusunnya dalam tabel. Prosesnya tiga tahap. Bayangkan kamu sedang menyiapkan kliping dari koran:

| Tahap | Yang terjadi | Analogi kliping |
|---|---|---|
| **1. Parsing** | Mengambil seluruh kode HTML halaman | Memfotokopi satu halaman koran |
| **2. Olah struktur HTML** | Memilih tag yang berisi data penting | Menggunting bagian yang dibutuhkan |
| **3. Framing** | Menyusun data menjadi tabel | Menempel guntingan di tabel yang rapi |

### TAHUKAH KAMU?

Tag HTML yang sering dicari: `h3` (judul), `p` (paragraf), `a` (tautan), dan `span` (potongan teks).

## Bagian 1 · Praktik web scraping

### Soal 2.1 DISKUSI Analogi versi kalian

Buat analogi baru untuk tiga tahap scraping dari kehidupan sehari-hari (bukan kliping koran).

| Tahap | Analogi kelompok kami |
|---|---|
| Parsing | |
| Olah struktur HTML | |
| Framing | |

LIVEWORKSHEETS

## Soal 2.2 PRAKTIK Scraping lowongan kerja

Buka notebook baru dan ketik kode berikut.

```
import requests
import pandas as pd
from bs4 import BeautifulSoup

th = 'https://www.jobs.id/lowongan-kerja?kata-kunci=part time'
halaman = requests.get(th)
hasil = BeautifulSoup(halaman.content, 'html.parser')
lowkers = hasil.find_all(class_='single-job-ads')

posisi = []
instansi = []
gaji = []

for p in lowkers:
    t1 = p.select('h3')
    t2 = t1[0].select('a')
    posisi.append(t2[0].get_text())

    t1 = p.select('p')
    t2 = t1[0].select('a')
    try:
        instansi.append(t2[0].get_text())
    except:
        instansi.append('-')

    t2 = t1[1].select('span')
    try:
        gaji.append(t2[1].get_text())
    except:
        gaji.append(t2[0].get_text())

lowker = pd.DataFrame({
    'Posisi': posisi,
    'Instansi': instansi,
    'Gaji': gaji
})
lowker
```

> ⚠ **PERHATIAN**
>
> Struktur website bisa berubah kapan saja, jadi hasilnya mungkin berbeda dari buku atau malah error. Kalau begitu, catat pesan error-nya dan tanyakan ke guru. Guru bisa mengarahkan ke website lain.

LIVEWORKSHEETS

| Pengamatan | Hasil |
|---|---|
| Jumlah baris lowongan pada tabel | |
| Kolom yang muncul | |
| Ada data kosong atau bertanda “-”? Apa penyebabnya? | |

### Soal 2.3 BACA KODE Apa yang dilakukan baris ini?

Isi kolom kanan dengan bahasa sendiri.

| Baris kode | Artinya |
|---|---|
| `halaman = requests.get(th)` | |
| `hasil.find_all(class_='single-job-ads')` | |
| `t1 = p.select('h3')` | |
| `instansi.append('-')` (di bagian except) | |
| `pd.DataFrame({...})` | |

### Soal 2.4 MISI Ganti topik

Ubah kata kunci `part time` pada alamat web menjadi `programmer`, jalankan lagi, dan bandingkan hasilnya.

| Pengamatan | Hasil |
|---|---|
| Jumlah lowongan sekarang | |
| Gaji tertinggi yang tampak | |
| Perubahan yang paling kalian sadari | |

LIVEWORKSHEETS

### Soal 2.5 CEK CEPAT Pilihan ganda

**1.** Hasil dari tahap parsing berupa …

**A** tabel rapi
**B** kode HTML halaman
**C** diagram batang
**D** kesimpulan

**2.** Mengapa kode scraping memakai try-except? …

**A** agar tampilan lebih indah
**B** supaya program tidak berhenti saat ada data yang tidak ditemukan
**C** agar lebih cepat
**D** untuk menghapus data

## Bagian 2 · Etika, privasi, dan keamanan data

**RINGKASAN MATERI**

Tidak semua data di internet boleh diambil sesukanya. Informasi dibedakan menjadi tiga jenis: **publik** (bisa dilihat siapa saja), **terbatas** (hanya untuk kelompok tertentu), dan **privat** (hanya untuk pemiliknya). Sebelum membuat scraper, tanyakan dulu: data ini milik siapa, dan apakah pemiliknya mengizinkan?

| Jenis data | Contoh | Boleh di-scrape? |
|---|---|---|
| Publik | Harga barang di toko online, judul lowongan | Umumnya boleh, tetap baca ketentuan situs |
| Terbatas | Postingan akun media sosial yang digembok | Tidak, kecuali ada izin |
| Privat | Isi email, pesan pribadi, data setelah login | Tidak |

LIVEWORKSHEETS

## Soal 2.6 DISKUSI Studi kasus

Putuskan boleh atau tidak, lalu beri alasan.

| Skenario | Boleh / Tidak | Alasan |
|---|---|---|
| Mengambil harga laptop dari lima toko online yang terbuka untuk umum, untuk dibandingkan | | |
| Mengambil daftar teman dan foto dari akun media sosial yang digembok | | |
| Mengambil nomor telepon dari profil orang lain untuk dikirimi iklan tanpa izin | | |

## Soal 2.7 DISKUSI Debat mini

Pernyataan: “Selama datanya terlihat di layar, boleh diambil bebas.” Setiap anggota memilih sikap, lalu kelompok menuliskan dua argumen dari masing-masing sisi.

| Setuju | Tidak setuju |
|---|---|
| 1. | 1. |
| 2. | 2. |

| Kesepakatan kelompok: |
|---|
| |

## Soal 2.8 CEK CEPAT Pilihan ganda

**1.** Manakah yang termasuk data privat?

A Harga barang di toko online

B Isi email setelah login

C Judul lowongan kerja

D Jadwal penerbangan

**2.** Langkah paling tepat sebelum melakukan scraping adalah …

A langsung mengambil semua data

B memastikan data bersifat publik dan membaca ketentuan situs

C menyembunyikan identitas

D mengabaikan aturan situs

LIVEWORKSHEETS

# 3

SUB BAB 3

# Visualisasi Data

*Rapikan data mentah, lalu ubah angka menjadi gambar yang bercerita.*

### RINGKASAN MATERI

Data hasil scraping masih mentah, jadi perlu **pra-pemrosesan** dulu: menghapus tanda yang mengganggu (misalnya titik pemisah ribuan), mengubah teks menjadi bilangan, dan mengganti nilai yang tidak bisa dihitung. Setelah rapi, data siap ditampilkan dalam diagram. Pilih jenis diagram sesuai tujuannya.

| Diagram | Cocok untuk | Contoh |
|---|---|---|
| **Batang** | Membandingkan nilai antar kategori | Gaji tiap lowongan |
| **Garis** | Melihat perubahan dari waktu ke waktu | Jumlah lowongan tiap bulan |
| **Lingkaran** | Melihat bagian terhadap keseluruhan | Proporsi lowongan per kota |
| **Titik (scatter)** | Melihat hubungan dua angka | Jam kerja dan gaji |

## Soal 3.1 JODOHKAN Pilih diagram yang tepat

Tulis jenis diagram dan alasan singkatnya.

| Kasus | Diagram | Alasan |
|---|---|---|
| Suhu kota Surabaya tiap jam selama sehari | | |
| Persentase siswa di tiap jenis ekstrakurikuler | | |
| Tinggi badan dan berat badan 30 siswa | | |
| Nilai ulangan rata-rata lima kelas | | |
| Jumlah pengunjung perpustakaan tiap bulan selama setahun | | |

LIVEWORKSHEETS

## Soal 3.2 PRAKTIK Merapikan data gaji

Ganti bagian pengambilan gaji pada kode Soal 2.2 dengan kode berikut, lalu jalankan.

```
    t2 = t1[1].select('span')
    try:
        xgaji = t2[1].get_text()
    except:
        xgaji = t2[0].get_text()
    xgaji = xgaji.replace('.', '')
    if xgaji == 'Gaji Dirahasiakan':
        xgaji = 0
    gaji.append(xgaji)
```

Apa fungsi `replace('.', '')`, dan mengapa "Gaji Dirahasiakan" diubah menjadi 0?

**Jawaban:**

## Soal 3.3 PRAKTIK Dari tabel jadi diagram

Tambahkan `import plotly.express as px` di baris paling atas, lalu ganti baris terakhir (`lowker`) dengan salah satu kode di bawah. Jalankan keduanya secara bergantian.

```
fig = px.bar(lowker, x='Posisi', y='Gaji')
fig.show()
```

```
fig = px.pie(lowker, values='Gaji', names='Posisi', title='Lowongan')
fig.show()
```

Menurut kelompok kalian, diagram mana yang lebih mudah dibaca untuk data lowongan ini? Mengapa?

**Jawaban:**

LIVEWORKSHEETS