

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI  
REPUBLIK INDONESIA  
2021

**Buku Panduan Guru  
Informatika untuk SMA Kelas X**

Penulis: Hanson Prihantoro Putro dkk.  
ISBN: 978-602-244-502-9

# Bab 6

## Analisis Data



*Gambar 6.1 Ilustrasi Unit Pembelajaran Analisis Data*

Sumber: <https://www.lego.com/en-us/aboutus/sustainability/people/responsible-business-principles>

Unit pembelajaran ini bertujuan untuk mengenalkan analisis data kepada siswa sehingga siswa dapat terbayang dengan proses yang terjadi dalam analisis data, khususnya secara otomatis. Siswa diajak untuk mempelajari IDE dan bahasa pemrograman untuk analisis data. Kemudian, siswa diajak untuk melakukan analisis data secara otomatis, mulai dari pengumpulan atau koleksi data otomatis melalui scraping website hingga visualisasi data ke dalam bentuk diagram.

## A. Tujuan Pembelajaran

Tujuan Pembelajaran untuk elemen Analisis Data di kelas X ialah siswa mampu:

1. Memahami bahwa data dapat dikoleksi melalui berbagai cara, baik manual maupun secara otomatis melalui perangkat.
2. Melakukan berbagai cara pengumpulan data yang dijelaskan, dan mengolah data yang dikumpulkan, dan menampilkannya.
3. Memahami aspek privasi dalam pengumpulan data.
4. Mengambil dan mempublikasi data dengan memperhatikan aspek privasi.
5. Memahami data yang terkumpul dalam jumlah besar dapat ditransformasi, digeneralisasi, disederhanakan untuk dimanfaatkan menjadi informasi.
6. Melakukan interpretasi data, menggali makna dan melakukan prediksi berdasarkan data yang ada.
7. Memahami pemodelan dan simulasi, dan kaitan data dengan model.
8. Melakukan penalaran dan prediksi berdasarkan model dan simulasi, dan memeriksa kesesuaian model terhadap data

## B. Kata Kunci

Analisis Data, Google Colaboratory, Koleksi Data, Web Scraping, Visualisasi Data

## C. Kaitan dengan Bidang PengetahuanLain

Analisis Data (AD) berkaitan dengan hampir semua elemen informatika terutama Algoritma dan Pemrograman, Teknologi Informasi dan Komunikasi, serta Dampak Sosial Informatika. Dalam melakukan analisis data, alat dan

teknologi informasi diperlukan agar data bisa diolah. Pengolahan ini juga akan lebih mudah jika dilakukan menggunakan bantuan pemrograman. Dari pengolahan data ini juga, aspek privasi data dan keamanan data perlu diperhatikan sebagai dampak sosial dari penerapan ilmu informatika.

## A. Strategi Pembelajaran

Mata pelajaran Informatika berkaitan erat dengan data. Data dikumpulkan dan digunakan untuk mengambil sebuah keputusan. Data yang dikumpulkan ini kemudian diolah dan ditampilkan dalam bentuk tekstual ataupun visual yang lebih mudah dimengerti. Untuk memulai mengolah data, unit pembelajaran Analisis Data akan memperkenalkan siswa tentang alat bantu (*tools*) yang akan digunakan untuk analisis data, yaitu IDE Google Colab dan bahasa pemrograman Python. Dari Google Colab, siswa akan belajar lingkungan pengembangan untuk membuat program yang akan melakukan analisis data. Dari Python, siswa akan belajar beberapa aturan dan fungsi sederhana yang akan digunakan untuk melakukan analisis data. Setelah mengenal kedua alat bantu tersebut, siswa mulai dapat membuat proyek analisis data, yaitu mengumpulkan/mengoleksi data melalui program web scraping yang dibuat dengan bahasa Python melalui IDE *Google Colab*. *Web scraping* ini mengambil data lowongan pekerjaan dari sebuah website, kemudian menampilkannya dalam bentuk tabel. Aktivitas dilanjutkan dengan mengolah data yang telah dikumpulkan dan ditampilkan dalam tabel tersebut (tekstual) menjadi tampilan diagram (visual).

Unit pembelajaran Analisis Data terdiri atas tiga subbab untuk tiga pertemuan dengan empat aktivitas ditambah satu pengantar, satu refleksi dan satu uji kompetensi. Sebelum memulai aktivitas pertama pada pertemuan pertama, siswa dapat diajak berdiskusi terkait kasus pencarian lowongan pekerjaan dari bagian Pengantar unit ini. Setelah itu, kegiatan dapat dilanjutkan dengan aktivitas pertama dan kedua, yaitu mengenal alat bantu Google Colab dan Python di pertemuan pertama Analisis Data. Kemudian, di pertemuan kedua, siswa diajak untuk melakukan aktivitas ketiga, yaitu membuat proyek *Web Scraping*. Selanjutnya, di pertemuan ketiga, siswa diajak untuk melakukan aktivitas keempat, yaitu proyek Visualisasi Data. Aktivitas

ini akan menunjukkan cara untuk memvisualisasikan data dalam salah satu bentuk diagram dan memperkenalkan diagram yang lain. Tidak lupa siswa di akhir pertemuan Analisis Data ini, siswa diajak untuk merefleksikan materi yang telah diperoleh sebelumnya melalui bagian Ayo, Renungkan yang dilanjutkan dengan bagian Uji Kompetensi. Diharapkan dengan mempelajari unit ini, utamanya siswa dapat memahami proses yang terjadi dalam analisis data khususnya koleksi data yang dilakukan secara otomatis melalui program atau perangkat lunak.

## D. Organisasi Pembelajaran

Tabel 6.1 Organisasi Pembelajaran Unit Analisis Data

Materi	Durasi (JP)	Tujuan Pembelajaran	Aktivitas
Mengenal Google Collab	1 JP	<ul style="list-style-type: none"> <li>Mengenal lingkungan pengembangan dari alat bantu untuk melakukan analisis data secara online.</li> <li>Menggunakan alat bantu untuk menulis, menjalankan dan mengembangkan program.</li> </ul>	Pertemuan-1: Penjelasan dengan demonstrasi, eksplorasi mandiri, diskusi
Mengenal Python	2 JP	<ul style="list-style-type: none"> <li>Mengenal bahasa dari alat bantu untuk melakukan analisis data secara online.</li> <li>Menggunakan alat bantu untuk melakukan pengolahan data secara sederhana.</li> </ul>	Pertemuan-1: Penjelasan dengan demonstrasi, eksplorasi, latihan <b>coding</b> , diskusi
Proyek <b>Web Scraping</b>	3 JP	<ul style="list-style-type: none"> <li>Memahami bahwa data dapat dikoleksi melalui berbagai cara, khususnya secara otomatis melalui perangkat, dalam hal ini perangkat lunak.</li> <li>Melakukan pengumpulan data sesuai cara yang dijelaskan, mengolah data yang dikumpulkan, dan menampilkannya.</li> <li>Memahami aspek privasi dan keamanan data.</li> </ul>	Pertemuan-2: Penjelasan dengan demonstrasi, latihan <b>coding</b> bersama, diskusi, refleksi tentang privasi dan keamanan data dalam <b>scraping</b> .

Materi	Durasi (JP)	Tujuan Pembelajaran	Aktivitas
Proyek Visualisasi Data	3 JP	<ul style="list-style-type: none"> <li>Memahami data yang terkumpul dalam jumlah besar dapat ditransformasi, digeneralisasi, disederhanakan untuk dimanfaatkan menjadi informasi.</li> <li>Melakukan interpretasi data dan menggali makna berdasarkan data yang ada.</li> </ul>	Pertemuan-3: Penjelasan dengan demonstrasi, latihan <b>coding</b> bersama, diskusi.

## E. Pengalaman Belajar Bermakna, Profil Pelajar Pancasila, Berpikir Komputasional dan Core Practice

Tabel 6.2 Pengalaman Bermakna, Profil Pelajar Pancasila, Praktik Inti, dan Berpikir Komputasional UnitAD

Pengalaman Bermakna	Profil Pelajar Pancasila	Berpikir Komputasional	Praktik Inti
Siswa memahami lingkungan tools analisis data	Mandiri	Abstraksi	Memahami <b>tools</b>
Siswa mengeksplorasi <b>tools</b> untuk analisis data berupa IDE dan bahasa pemrograman	Mandiri, Bernalar Kritis	Abstraksi, Algoritma, Dekomposisi, dan Pengenalan Pola	Mengimplementasikan <b>tools</b> yang sesuai
Siswa berkolaborasi dan berdiskusi mengembangkan kasus Scraping Website	Gotong Royong, Bernalar Kritis, Kreatif	Abstraksi, Algoritma, Dekomposisi, dan Pengenalan Pola	Menyelesaikan kasus pemrograman visual
Siswa berkolaborasi dan berdiskusi mencoba melakukan pengembangan analisis data, secara unplugged maupun plugged.	Gotong Royong, Bernalar Kritis, Kreatif	Abstraksi, Algoritma, Dekomposisi, dan Pengenalan pola	Mengimplementasikan pemrograman manual

## F. Panduan Pembelajaran

Pembelajaran pada unit pembelajaran Analisis Data ini dapat dilakukan mulai dari pengenalan IDE (Google Colab), pengenalan bahasa (Python) yang dilanjutkan dengan pembuatan Proyek *Web Scraping*. Kemudian, siswa dapat

diajak untuk melanjutkan analisis data ini ke aktivitas pembuatan Proyek Visualisasi Data dan *Web Scraping* Lanjutan.

## 1. Pertemuan 1: Mengenal Google Colab (1 JP)

### Tujuan Pembelajaran:

1. Mengetahui lingkungan pengembangan dari alat bantu untuk melakukan analisis data secara *online*.
2. Menggunakan alat bantu untuk menulis, menjalankan, dan mengembangkan program.

### Apersepsi dan Pemantik

Siswa diajak untuk mengingat kembali alat-alat apa saja yang digunakan saat belajar modul Algoritma Pemrograman. Salah satu yang menjadi penekanan, yaitu Code::Block sebagai IDE (lingkungan pengembangan terintegrasi) yang digunakan untuk membuat program C. Kali ini, siswa diajak untuk mempelajari salah satu IDE lain yang sering digunakan dalam analisis data secara *online*, yaitu Google Colab (Aktivitas AD-K10-01-P Mengenal Google Colab).

### Kebutuhan Sarana dan Prasarana

1. Komputer dan Internet
2. Google Colab yang diakses di <https://colab.research.google.com/>

### Aktivitas

Guru dapat menampilkan Google Colab di komputernya yang kemudian ditampilkan ke siswa melalui proyektor. Kemudian, siswa diminta membuka Google Colab di komputer masing-masing. Siswa dipandu untuk membuat sebuah file atau *notebook* baru. Biasanya, file baru dinamai “Untitled0.ipynb”. Akan lebih bagus jika siswa sudah memiliki akun Google/GMail. Setelah tampil sebuah file atau *notebook* kosong, siswa dapat diajak untuk mengenali ada objek apa saja di area kerja IDE Google Colab. Dengan identifikasi ini, kemampuan siswa dalam hal Berpikir Komputasional (BK) khususnya dalam aspek dekomposisi dapat digali di sini. Selain itu, kemampuan BK dalam aspek pengenalan pola dapat diasah dengan

membandingkan setiap objek tersebut dengan IDE lain seperti Code::Block. Beberapa objek yang penting ialah Nama *Notebook* dan Konten *Notebook*.

Setelah mengenali objek-objek dalam area kerja Google Colab, siswa dapat diajak untuk mulai menulis program yang sangat sederhana, yaitu program “Hello World”. Program ini ialah program yang hampir selalu disediakan setiap pertama kali mempelajari bahasa pemrograman, termasuk IDE-nya. Selain teks “Hello World”, guru juga dapat menuliskan teks lain misal “Halo Dunia” atau yang lain. Setelah itu, siswa diajak untuk bersama-sama mengeksekusi kode program tersebut. Istilahnya ialah *running program* atau menjalankan program. Pastikan bahwa semua siswa dapat menampilkan tulisan “Hello World” atau sesuai dengan yang diketik sebelum berlanjut ke aktivitas berikutnya.

## 2. Pertemuan 1 : Mengenal Python (2 jp)

### Tujuan Pembelajaran:

1. Mengenal bahasa dari alat bantu untuk melakukan analisis data secara *online*.
2. Menggunakan alat bantu untuk melakukan pengolahan data secara sederhana.

### Apersepsi dan Pemantik

Siswa diajak untuk mengingat kembali bahasa pemrograman apa yang dipelajari pada modul Algoritma Pemrograman, dan mempelajari ulang tabel terjemahan bahasa C ke bahasa Python yang sudah diberikan. Kali ini, siswa diajak untuk mencoba memrogram dalam bahasa Python sebagai sebuah bahasa yang sering digunakan dalam analisis data.

### Kebutuhan Sarana dan Prasarana

1. Komputer dan Internet
2. Google Colab yang diakses di <https://colab.research.google.com/>

### Aktivitas

Kegiatan ini ialah kegiatan lanjutan setelah siswa mengenal Google Colab pada aktivitas **AD-K10-01-P** (Mengenal Google Colab). Siswa dapat diajak

berkenalan dengan Python dengan melakukan aktivitas **AD-K10-02-P** (Mengenal Python). Siswa diajak belajar mulai dari penulisan, penyimpanan data, perulangan, percabangan dan penggunaan *library*. Pembelajaran pemrograman di modul Analisis Data hanya akan belajar program Python secara sederhana. Siswa dapat disemangati bahwa di sini tidak akan belajar bahasa Python yang kompleks dan rumit karena inti pembelajarannya ada di aspek analisis datanya. Siswa hanya akan belajar beberapa aturan penulisan yang nantinya akan digunakan untuk membuat proyek *Web Scraping*.

Proses pembelajaran dilakukan secara eksploratif untuk setiap konsep dan aturan penulisan. Siswa tidak akan diberikan permasalahan pemrograman seperti di unit Algoritma Pemrograman. Pada unit ini, siswa hanya diminta untuk memahami hasil eksekusi dari setiap contoh program Python yang diberikan. Siswa lebih belajar tentang pembacaan kode (*programreading and comprehension*), serta *coding* dari contoh, *bukan* mengonstruksi program (*program construction*). Dalam konstruksi program, siswa menyelesaikan persoalan dan memikirkan kode solusinya. Kemampuan untuk memakai kode yang sudah ada untuk menyelesaikan solusi sering kali dibutuhkan di masa depan karena makin banyak kode yang tersedia.

Berikut beberapa jawaban dari latihan yang diberikan. Jawaban-jawaban ini juga dapat diperoleh jika siswa mencobanya secara langsung di Google Colab masing-masing dan guru juga bisa menunjukkannya di Google Colab komputer guru.

### *Jawaban Latihan 1*

1. **Print("2+3")** → Teks "2+3" karena dua ditambah tiga ditulis adanya sebagai sebuah teks dalam tanda petik (string). Tidak ada bilangan dan tidak ada operasi penjumlahan.
2. **Print(2+3)** → Bilangan 5 karena dua dan tiga tanpa tanda petik diterjemahkan sebagai bilangan, begitu juga tanda plusnya. Maka, perintah ini mencetak hasil dari penjumlahan bilangan dua dengan bilangan tiga.
3. **Print("2"+"3")** → Teks "23", karena dua dan tiga dengan tanda petik diterjemahkan sebagai teks (string), di mana penjumlahan teks dilakukan dengan menjejerkan kedua teks tersebut (*concatenation*).

4. Penggunaan tanda petik bermakna bahwa apapun yang berada di antara tanda petik diterjemahkan sebagai teks atau string. Jika dicetak, dicetak apa adanya. Jika dioperasikan, mengikuti operasi teks.

#### *Jawaban Latihan 2*

1. Terdapat tiga variabel, yaitu **bil1**, **bil2** dan **jumlah**.
2. Makna dari baris ke-2 (**bil2 = 5**) ialah variabel **bil2** diisi dengan bilangan 5.
3. Makna dari baris ke-3 ialah variabel **jumlah** diisi dengan hasil penjumlahan dari isi variabel **bil1** dan **bil2**, yaitu 10+5 atau 15.

#### *Jawaban Latihan 3*

1. Variabel **bil2** tidak tercetak di hasil eksekusi karena perintah mencetak nama variabel hanya dapat dilakukan pada perintah di baris terakhir (dalam hal ini mencetak isi variabel jumlah).
2. Adapun dengan menggunakan perintah **print**, kita dapat mencetak nilai dari variabel di manapun. Dengan demikian, menggunakan **print**, kedua variabel dapat tercetak.

#### *Jawaban Latihan 4*

1. Output dari variabel **bil1** pada baris keempat ialah 5 saja.
2. Nilai 10 yang diisi sebelumnya hilang, ditimpa oleh 5. Pada saat variabel **bil1** diisikan (*assigned*) dengan 5 pada baris ketiga, isi dari variabel ini diganti dengan 5. Apapun nilai sebelumnya akan diganti/hilang. Inilah sifat pemrograman sekuensial di mana isi dari sebuah variabel bergantung pada nilai terakhir yang diisikan.

#### *Jawaban Latihan 5*

1. Yang tercetak di output ialah bilangan 9.
2. Hal ini terjadi karena array dalam Python dimulai dari indeks 0. Maka, **print(data[1])** maknanya ialah mencetak elemen ke-1 atau urutan kedua. Jika ingin mencetak bilangan 10, yang ditulis ialah **print(data[0])**.

### *Jawaban Latihan 6*

1. Perlu ditambah 1 agar berubah/berpindah ke indeks berikutnya. Jika baris kelima ini dihapus, indeks akan terus berada di nilai 0, tidak sesuai dengan elemen yang dicetak.
2. Jika diganti dengan 2, perubahan antarindeks terjadi dalam selisih 2 seperti 0, 2, 4, dan seterusnya, tidak sesuai dengan elemen yang dicetak.
3. Jika dipindah ke baris keempat, yang tercetak dimulai dari 1, 2, 3, dan seterusnya, tidak sesuai dengan elemen yang dicetak.

### *Jawaban Latihan 7*

Struktur loop while dalam Python

```
while <kondisi_perulangan>:  
    <instruksi_perulangan>
```

Siswa dapat belajar lebih banyak tentang loop while dari internet <https://realpython.com/python-while-loop/> atau <https://www.petanikode.com/python-perulangan/>

### *Jawaban Latihan 8*

1. Pesan *error* yang muncul ialah "**unsupported operand type(s) for /: 'str' or 'int'**"
2. Maksudnya operan yang diberikan tidak bisa dilakukan operasi dengan operator pembagian (/). Hal ini terjadi karena data pada elemen ke-2 bernilai "Cindi" yang berupa nilai teks atau string. Nilai teks ini tidak bisa dibagi. Oleh karena itu, program memberikan pesan kesalahan.

### *Jawaban Latihan 9*

Bilangan yang tercetak ada tiga, yaitu 5, 4.5 dan 2. Adapun data yang lain bukan bilangan yang bisa dibagi.

### *Jawaban Latihan 10*

Struktur branch if dalam Python

```
if <kondisi_percabangan>:  
    <instruksi_saat_kondisi_benar>  
else:  
    <instruksi_saat_kondisi_salah>
```

Siswa dapat belajar lebih banyak tentang loop while dari internet <https://realpython.com/python-conditional-statements/> atau <https://www.petanikode.com/python-percabangan/>

### 3. Pertemuan 2: Proyek Web Scraping (3 jp)

#### Tujuan Pembelajaran:

1. Memahami bahwa data dapat dikoleksi melalui berbagai cara, khususnya secara otomatis melalui perangkat, dalam hal ini perangkat lunak.
2. Melakukan pengumpulan data sesuai cara yang dijelaskan, dan mengolah data yang dikumpulkan, dan menampilkannya.
3. Memahami aspek privasi dan keamanan data.

#### Apersepsi dan Pemantik

Siswa dapat diajak mengingat kembali latar belakang dari unit Analisis Data ini yang berada di Buku Siswa bagian Penjelasan terkait kepentingan sebuah analisis data dan proses *web scraping*.

#### Kebutuhan Sarana dan Prasarana

1. Komputer dan internet
2. Google Colab yang diakses di <https://colab.research.google.com/>

#### Aktivitas

Kegiatan ini ialah kegiatan utama di unit Analisis Data, yaitu tentang koleksi data. Setelah mengenal Google Colab dan Python, siswa diajak membuat proyek kecil untuk mengoleksi data dari sebuah *website* lowongan pekerjaan melalui aktivitas **AD-K10-03-P** (Proyek Web Scraping). Langkah pengoleksian data dijelaskan secara umum melalui alur proses *web scraping* yang diperlihatkan melalui Gambar 6.13 Buku Siswa. Guru dapat menjelaskan alur ini terlebih dahulu sebelum masuk ke detail setiap kegiatan.

Di akhir proyek, siswa diajak untuk membandingkan hasil *web scraping* ini dengan halaman web lowongan pekerjaan yang asli. Diperlihatkan bahwa data lowongan pekerjaan di web aslinya dengan berbagai informasi yang mungkin cenderung berlebihan, di program *web scraping* ini data diperlihatkan secara

lebih ringkas. Dari ringkasan hasil koleksi data ini, proses pengolahan atau analisis data dapat dilanjutkan secara lebih mudah. Hal ini menjadi salah satu keuntungan dari proses koleksi untuk meringkas data (*summarize*).

#### 4. Pertemuan 3: Proyek Visualisasi Data (3 JP)

##### Tujuan Pembelajaran:

1. Memahami data yang terkumpul dalam jumlah besar dapat ditransformasi, digeneralisasi, disederhanakan untuk dimanfaatkan menjadi informasi.
2. Melakukan interpretasi data dan menggali makna berdasarkan data yang ada.

##### Apersepsi dan Pemantik

Siswa dapat diajak mengingat kembali latar hasil dari Aktivitas 3 yang berupa tabel. Siswa diminta membandingkan yang ditampilkan di tabel dengan yang ditampilkan di halaman web lowongan pekerjaan, mana yang lebih mudah dipahami. Di sinilah, analisis data bekerja, mengolah data agar menjadi lebih bermakna. Namun, tidak sampai di sini. Kita dapat mengolah kembali data tabel tersebut agar lebih dapat dipahami lagi, yaitu dengan menampilkannya dalam bentuk visual atau diagram.

##### Kebutuhan Sarana dan Prasarana

1. Komputer dan internet
2. Google Colab yang diakses di <https://colab.research.google.com/>

##### Aktivitas

Kegiatan ini merupakan kegiatan lanjutan dari kegiatan utama koleksi data. Setelah data dikoleksi, data kemudian ditampilkan dengan tampilan yang lebih informatif dalam bentuk grafik/diagram/*chart*. Siswa diajak untuk membuat program yang akan mengolah data tekstual menjadi data visual. Data visual ini diperlukan untuk menunjukkan hubungan antardata misalnya posisi lowongan pekerjaan dengan gaji yang ditawarkan. Aktivitas ini bisa dilakukan oleh siswa pada aktivitas **AD-K10-04-P** (Proyek Visualisasi Data).

Sebelum visualisasi dibuat, terlebih dahulu data yang diperoleh di aktivitas sebelumnya (*web scraping*) ditata dalam langkah pra-pemrosesan (*pre-processing*). Penataan dilakukan agar visualisasi dengan beberapa perhitungan di langkah

berikutnya dapat dilakukan. Dalam proyek ini, siswa diajak untuk memvisualisasikan dalam bentuk diagram batang. Setelah itu, siswa diajak untuk melakukan eksplorasi terkait pembuatan diagram yang lain misalnya diagram lingkaran.

### *Jawaban Latihan 11*

Membuat diagram scatter dari data lowongan pekerjaan dapat dilakukan dengan perintah

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

Untuk penggunaan diagram, kasus lain yang lebih cocok menggunakan diagram lingkaran, misalnya persebarang lowongan pekerjaan di setiap lokasi/kota, berapa persen lowongan di Kota A, berapa persen di Kota B. Adapun kasus lain yang lebih cocok untuk penggunaan diagram *scatter* misalnya jumlah lowongan pekerjaan dari tahun ke tahun.

## **G. Pengayaan Aktivitas Utama**

Pengayaan dapat dilakukan dengan mengajak siswa untuk merefleksikan kegiatan pembelajaran terkait penggunaan bahasa pemrograman, target *scraping*, sistem visualisasi *real-time* serta pemanfaatan analisis data misalnya untuk keperluan prediksi.

Terkait penggunaan bahasa pemrograman, siswa dapat diajak untuk membandingkan kegiatan pemrograman di unit Algoritma Pemrograman dengan yang ada di unit Analisis Data ini. Pada unit ini, siswa lebih diajak untuk memanfaatkan pemrograman sebagai alat untuk melakukan Analisis Data. Pembelajaran di sini dilakukan dengan mode eksplorasi di mana siswa diminta untuk mengikuti dan memahami kode program yang diberikan bukan untuk mengonstruksi kode program dari nol untuk melakukan *problem solving* sesuai kasus yang diberikan seperti pada unit Algoritma Pemrograman. Bahasa yang digunakan juga relatif lebih sederhana, hanya beberapa sintaks yang digunakan dalam proyek *web scraping* dan visualisasi data.

Dari mesin *scraper* yang dibuat, program ini sebenarnya serupa dengan mesin pencarian yang sering digunakan siswa. Kedua mesin sama-sama melakukan penelusuran pada setiap halaman web yang ditemukan, mengambil data yang diperlukan dan menampilkannya ke layar. Bedanya ialah mesin

*scraper* yang dikerjakan dalam proyek hanya melakukan *scraping* satu halaman website untuk dianalisis, sedangkan mesin pencarian melakukan *scraping* pada semua halaman web yang ada di internet. Tentunya, penggunaan indeks yang menjadi acuan, memudahkan mesin pencari untuk hanya *scraping* pada halaman yang relevan saja. Kemudian, dalam kasus proyek yang dikerjakan, mesin *scraper* mengambil data lowongan pekerjaan berupa posisi lowongan, instansi pemberi lowongan, dan gaji yang diberikan. Adapun pada mesin pencari, data yang diambil ialah judul halaman web beserta beberapa isi konten halaman yang relevan dengan kata kunci yang digunakan.

Kemudian, terkait target *scraping* pada proyek *web scraping*, siswa melakukan *scraping* pada halaman public, yaitu halaman yang bisa diakses siapa saja yang memiliki alamat tersebut. Namun demikian, sebenarnya, *scraping* juga bisa dilakukan pada halaman privat seperti email atau media sosial seseorang. Untuk bisa melakukannya, mesin *scraper* perlu memegang izin (*authorization*) untuk bisa masuk ke halaman privat tersebut. Bagi *web email* ataupun media sosial yang memang membolehkan ini, mereka akan memberikan fasilitas bagi program lain untuk bisa mengakses halaman di dalam sistem mereka. Seperti halnya kita yang perlu *username* dan *password* untuk bisa masuk ke email kita, program lain yang akan masuk ke dalam sistem email (seperti *scraper*) juga memerlukan semacam *username* dan *password* yang menjadi pertanda bahwa program tersebut memang diizinkan untuk masuk ke dalam sistem. Tentunya konten yang bisa diakses terbatas kepada akun yang memberikan izin untuk program tersebut bisa masuk ke dalam sistem.

Selanjutnya, terkait alat analisis atau *tools* yang digunakan, *library* yang disediakan oleh Python membantu kita untuk melakukan *scraping* sehingga beberapa proses dapat dilakukan hanya dengan satu baris kode program. Padahal sebenarnya, *library* Python telah menyembunyikan kerumitannya. *Library BeautifulSoup* misalnya yang digunakan untuk melakukan *parsing*, mengambil kode program HTML yang menyusun sebuah halaman website. Sebenarnya, cara yang dilakukan tidak jauh berbeda dengan saat kita membuka sebuah halaman pada *browser*. Namun demikian, browser menginterpretasi setiap HTML menjadi sebuah halaman website sesuai dengan kode dan tag-

nya. Adapun parser BeautifulSoup mengambil setiap kode HTML apa adanya. Dari sini, kemudian program *webscraper* siswa mengolahnya sehingga hanya beberapa kode dan data yang relevan dengan kasus saja yang ditampilkan. Selain BeautifulSoup, Python memiliki banyak library lain seperti *Pandas* untuk melakukan pengolahan data, *Plotly* untuk membuat diagram atau chart, *NumPy* untuk melakukan perhitungan ilmiah, *Matplotlib* untuk melakukan perhitungan dan visualisasi matematis, dan *library* lainnya yang banyak digunakan dalam proses-proses analisis data.

Refleksi selanjutnya pada sistem visualisasi yang dilakukan secara *real-time*, bahwa pada proyek *web scraping* yang dilakukan siswa, program hanya akan menampilkan data di *website* lowongan pekerjaan saat kita melakukan klik “Run”. Setelah data tampil, dan kemudian *website* mengalami perubahan, tampilan dari *scraper* yang dibuat siswa akan tetap dan tidak ikut berubah. Ini artinya, program *web scraping* ini tidak dibuat secara *real-time*. Akan sangat menarik jika data lowongan pekerjaan di program *scraper* juga ikut berubah, terbaharui sesuai perubahan yang terjadi di web aslinya. Ini yang disebut sistem *real-time*, seperti halnya tampilan lokasi ojek saat kita melakukan pemesanan online, tampilan informasi kedatangan atau keberangkatan pesawat, atau kondisi kurs mata uang di bank.

Ilustrasi kerja sama dapat diperlihatkan dari ketiga sistem tersebut. Misalnya, pada sistem ojek online, sistem GPS akan mengirimkan data koordinat lokasi dari pengendara/*driver*. Data ini kemudian digunakan untuk memperbaharui lokasi pengemudi ojek (*driver*) yang kita pesan di peta. Setiap waktu, data ini dikirim sehingga *driver* akan terlihat bergerak juga setiap waktu. Namun demikian, kadang *driver* terlihat bergerak tidak lancar. Hal ini bisa terjadi karena gangguan jaringan yang tidak lancar. Mungkin ada beberapa waktu di mana data lokasi *driver* tidak terkirim yang mengakibatkan gambar *driver* di peta juga tidak terlihat berubah atau berubah, tetapi tidak lancar.

Terakhir, refleksi tentang prediksi. Jika mencari menggunakan mesin pencari Google, kita dapatkan bahwa lowongan pekerjaan programmer di setiap tahun diperoleh data sebagai berikut.

No	Tahun	Hasil Pencarian Lowongan
1	2018	185.000
2	2019	426.000
3	2020	438.000

Dari sini, diperoleh informasi bahwa dari tahun 2018 ke 2019 hingga tahun 2020, jumlah lowongan pekerjaan selalu bertambah. Dengan demikian, dapat kita prediksi atau perkirakan, bahwa lowongan pekerjaan di tahun 2021 juga akan bertambah, mungkin sekitar 450.000 lowongan.

Aktivitas pembelajaran lain bisa dikembangkan dengan mempelajari materi dari situs-situs yang juga sempat dibahas di materi utama, termasuk untuk jawaban penugasannya:

1. Jobs ID (2020). Info Lowongan Kerja Terbaru dan Populer 2020. Diakses dari <http://jobs.id>
2. Wikipedia (2020). Web Scraping. Diakses dari [https://en.wikipedia.org/wiki/Web\\_scraping](https://en.wikipedia.org/wiki/Web_scraping)
3. Google Colaboratory (2020). Welcome to Colaboratory. Diakses dari <https://colab.research.google.com/>
4. Dataquest (2020). Tutorial: Web Scraping with Python using BeautifulSoup. Diakses dari <https://www.dataquest.io/blog/web-scraping-tutorial-python/>
5. PyData (2021). Pandas: Python Data Analysis Library. Diakses dari <https://pandas.pydata.org/>
6. RealPython (2021). Python While Loops (Indefinite Iteration). Diakses dari <https://realpython.com/python-while-loop/>
7. PetaniKode(2021).BelajarPemrogramanPython:MemahamiPerulangan. Diakses dari <https://www.petanikode.com/python-perulangan/>

8. RealPython (2021). Condition Statement in Python. Diakses dari <https://realpython.com/python-conditional-statements/>
9. Petani Kode (2021). Belajar Pemrograman Python: Memahami Percabangan untuk Membuat Logika Pemrograman. Diakses dari <https://www.petanikode.com/python-percabangan/>

## H. Asesmen dan Rubrik Penilaian

*Tabel 6.3. Rubrik Penilaian Unit Pembelajaran Analisis Data*

Nomor Soal	Indikator Capaian	Baik Sekali	Baik	Cukup	Kurang
1	Memahami cara-cara koleksi data	Penjelasan terkait langkah koleksi data benar dan mudah dipahami.	Penjelasan terkait langkah koleksi data benar, tetapi sulit dipahami.	Penjelasan terkait langkah koleksi data masih terdapat kesalahan minor.	Penjelasan terkait langkah koleksi data terdapat banyak kesalahan mayor.
2	Mengenal alat bantu untuk melakukan analisis data (Python)	Dapat menunjukkan kode program yang bersesuaian dan memodifikasinya dengan tepat.	Dapat menunjukkan kode program yang bersesuaian dan belum bisa memodifikasinya.	Tidak dapat menunjukkan kode program yang diperlukan, tetapi dapat menjelaskan ide perubahannya.	Tidak dapat menjelaskan ide perubahan kode program untuk menyelesaikan persoalan.
3	Melakukan interpretasi dan menggali makna berdasarkan data yang ada	Data berhasil diperlihatkan sesuai dengan yang diharapkan.	Data dapat dianalisis dan dibuat programnya, tetapi belum berhasil memperlihatkan yang diinginkan.	Data yang bersesuaian dapat ditunjukkan, tetapi belum bisa membuat kodenya.	Tidak dapat menunjukkan data yang diperlukan.

Nomor Soal	Indikator Capaian	Baik Sekali	Baik	Cukup	Kurang
4	Membuat rangkuman data sesuai spesifikasi dan tujuan	Hasil rangkuman (baik dari koran atau website lain) benar dan dapat ditampilkan dengan baik.	Hasil rangkuman benar, tetapi dengan tampilan yang kurang baik.	Hasil rangkuman masih terdapat kesalahan minor.	Hasil rangkuman terdapat banyak kesalahan mayor.
5	Melakukan berbagai cara pengumpulan data	Penjelasan terkait proses perangkuman yang dilakukan benar dan mudah dipahami.	Penjelasan terkait langkah perangkuman yang dilakukan benar, tetapi sulit dipahami.	Penjelasan terkait langkah perangkuman data masih terdapat kesalahan minor.	Penjelasan terkait langkah perangkuman data terdapat banyak kesalahan mayor.

## I. Jawaban Uji Kompetensi

### Soal 1

Jawaban siswa dapat beragam, tetapi semuanya harus dalam bingkai alur seperti diperlihatkan pada Gambar 6.13 Buku Siswa, yaitu mulai dari melakukan *parsing*, memproses struktur HTMLnya hingga menampilkannya atau *framing* dalam tabel.

### Soal 2

Mengubah agar program menampilkan lowongan pekerjaan lain berarti mengubah alamat yang digunakan sebagai sumber web yang akan dilakukan parsing. Dalam Buku Siswa, diperlihatkan pada Gambar 6.14 Baris 5. Di baris tersebut terdapat alamat web `jobs.id` yang diberi kata kunci “part time” sebagai pekerjaan yang dicari. Maka, jika ingin mencari pekerjaan “programmer”, baris tersebut dapat diganti sebagai berikut.

5	<code>web = "https://www.jobs.id/lowongan-kerja?kata-kunci=programmer"</code>
---	---

### Soal 3

Untuk melihat data lokasi di dalam lowongan pekerjaan, kita perlu melihat struktur HTML website tersebut. Hal ini dapat diperoleh setelah proses parsing pada Gambar 6.14 Buku Siswa dilakukan. Dari proses tersebut, kita peroleh kode sebagai berikut.

```
1 <div class="col-xs-12 single-job-ads">
2 <div class="col-xs-12 col-md-10">
3 <div class="col-xs-2 col-md-2">
4 
5 </div>
6 <div class="col-xs-8 col-md-10">
7 <h3>
8 <a class="bold" href="https://www.jobs.id/lowongan/
  MTc2Nzk4/part-time-corporate-investment-consultant-
  indosukses-futures-pt?qt_ref=search&qt_
  page=1&qt_pos=0" target="_blank">Part Time -
  Corporate Investment Consultant</a>
9 </h3>
10 <p>
11 <a class="bold" href="https://www.jobs.id/lowongan-
  kerja-indosukses-futures-pt?kata-kunci=Indosukses+Futur
  es+PT">Indosukses Futures PT</a>
  -
12 <i class="fa fa-map-marker"></i>
13 <span class="location">Jakarta Barat</span>
14 </p>
15 <p>
16 <span class="semi-bold currency text-success">IDR </span>
17 <span class="semi-bold">1.000.000</span>
18 -
19 <span class="semi-bold">2.000.000</span>
20 </p>
21 <p>
22 Bertanggung jawab untuk prospek klien. Menyediakan review
23 dan saran finansial utk klien.
  </p>
24 <p class="text-muted">
25 1 minggu yang lalu
26 </p>
27 </div>
28 </div>
29 </div>
30
```

Jika diperhatikan, data lokasi lowongan berada satu hierarki dengan informasi institusi, yaitu sama-sama berada di dalam tag `<p>`. Menariknya, pada informasi ini, diberikan penanda atribut kelas “location” sehingga memudahkan kita untuk mengidentifikasinya. Namun sebelumnya, kita perlu mempersiapkan array untuk menampung data lokasi di tempat deklarasi yang sama dengan array yang lain, yaitu di Gambar 6.15 Baris 13.

```
13 lokasi = []
```

Kemudian setelah itu, data lokasi bisa diambil di variabel lowongan pekerjaan yang telah diiterasi dalam **for**, yaitu variabel **p**. Dari variabel ini, kita ambil data yang mengandung “location”. Setelah itu, data ini disisipkan ke dalam array lokasi.

```
26
27 t2 = p.select(".location")
28 lokasi.append(t2[0].get_text())
```

Setelah itu, array lokasi ini dapat ditampilkan saat membingkai datanya ke dalam tabel, sebagai mana kita menambahkan Baris 39 ke dalam Gambar 6.16 pada Buku Siswa seperti berikut.

```
35
36 lowker = pd.DataFrame({
37     "Posisi": posisi,
38     "Instansi": instansi,
39     "Lokasi": lokasi,
40     "Gaji": gaji
41 })
42 Lowker
```

#### Soal 4

Kegiatan ini dapat diberikan sebagai alternatif (aktivitas *unplugged*) jika tidak dimungkinkan melakukan *web scraping* secara *online* menggunakan komputer. Siswa diminta untuk mencari di media cetak seperti koran atau majalah, kemudian menemukan daftar lowongan pekerjaan di media tersebut. Selanjutnya, siswa diminta untuk merangkum lowongan pekerjaan yang ada dalam sebuah tabel yang berisi Posisi/Jabatan, Nama Perusahaan/Instansi serta Gaji yang ditawarkan jika ada. Kemudian, siswa diminta untuk merefleksikan aktivitas apa saja yang mereka lakukan terkait pengumpulan data lowongan pekerjaan dari media cetak tersebut.

### Soal 5

Kegiatan ini diberikan sebagai pengembangan dari kegiatan proyek *web scraping* untuk kelas yang memiliki kemampuan lebih. Setelah diminta untuk mencari berbagai *website* terkait lowongan pekerjaan, siswa diminta untuk melihat struktur kode HTML dari *website* yang mereka temukan. Dari sekian *website* yang ada, banyak diantaranya yang memiliki struktur yang rumit sehingga susah untuk dilakukan analisis dan *scraping*. Namun demikian, masih ada beberapa *website* yang memiliki kode HTML yang sederhana sehingga relatif lebih mudah untuk dianalisis strukturnya dan dilakukan *scraping*. Salah satunya ialah *website* JobsDB.

Siswa diminta untuk melakukan apa yang sudah dilakukan di Aktivitas 3 dengan mengganti alamat *website* nya. Dengan mengganti *website*, tentunya proses *scraping* di program Python juga berubah. Siswa diminta untuk mengubah kode program *web scraping* sebelumnya sehingga dapat melakukan *scraping* dari *website* JobsDB. Aktivitas ini dan menghasilkan daftar lowongan pekerjaan dalam bentuk tabel. Dalam aktivitas ini, siswa diberikan tabel daftar periksa (*checklist*) yang memandu aktivitas siswa dalam membuat program *web scraping* menggunakan bahasa Python.

### J. Interaksi Guru dan Orang Tua/Wali

Orang tua/wali hendaknya selalu aktif dalam mengawasi anaknya ketika melakukan aktivitas khususnya dalam penggunaan dan eksplorasi internet. Guru dapat berinteraksi dengan memberikan informasi dan tips bagi orang tua dalam baik dalam penguasaan konsep dan *tools*. Di rumah, orang tua/wali dapat membantu menjelaskan terkait proses koleksi data yang dapat mereka lakukan di masa depan, khususnya dalam melakukan pencarian lowongan pekerjaan.

### Kegiatan Refleksi Bersama

Setelah melalui beberapa bab pembelajaran, guru dapat melakukan refleksi bersama tim pengajar (jika ada). Guru juga dapat melibatkan siswa untuk menggali minat mereka terhadap topik pembelajaran atau aktivitas yang telah dilakukan. Berikut adalah contoh pertanyaan refleksi yang bisa dilakukan bersama, baik bersama tim guru maupun siswa.

1. Berdasarkan bab/aktivitas pembelajaran yang sudah berlangsung, bab/aktivitas mana yang memiliki respon paling positif dan respon paling negatif di kelas? Mengapa?
2. Bab/aktivitas mana yang paling anda kuasai?
3. Pada bab/aktivitas apa anda merasa kreatif ketika mengajar?
4. Jika anda harus bercerita kepada rekan sesama guru mengenai hal yang telah anda pelajari selama mengajar, apa yang akan anda ceritakan?
5. Jika anda dapat bertanya kepada tim guru informatika di sekolah/kelompok kerja guru informatika/komunitas guru informatika mengenai pembelajaran informatika, hal apa yang ingin anda tanyakan?
6. Jika anda memiliki kesempatan untuk mengulang bab/topik yang sudah berlangsung, topik apa yang menurut anda perlu disampaikan lagi dengan cara yang berbeda?