

Eksplorasi Pola Penjualan Produk di *E-Commerce* menggunakan *Web Scraping*

Deigo Anugrah Pratama¹, Kevin Rafael Manuela Feltya², Ibnu Rizal Mutaqin Muyasar³,
Belsana Butar Butar⁴, Elly Indrayuni⁵

Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Indonesia
E-mail: deigoanugrahp@gmail.com¹, kevinrafaelmanuelafeltya@gmail.com²,
ibnumuyasar12@gmail.com³, belsana.bbb@bsi.ac.id⁴, elly.eiy@bsi.ac.id⁵

Article History:

Received: 31 Oktober 2025

Revised: 01 Desember 2025

Accepted: 10 Desember 2025

Keywords: *Web Scraping, Shopee, Google Collaboratory, Margin of Error, Data Visualization, Looker Studio, Review Data.*

Abstract: *This study examines product sales patterns and review activity data on the Shopee e-commerce platform by employing web scraping techniques on customer review data as a proxy for transaction dates. The objectives of the research include exploring product sales patterns, identifying time-based sales spikes, and presenting interactive visualizations as decision-support tools. Data were collected through JavaScript-based web scraping executed in the browser's developer console to extract customer reviews. The data were then processed during the preprocessing stage using Google Colab and Python libraries, and subsequently uploaded to Looker Studio to create interactive dashboards. The visualizations produced in Looker Studio support the optimization of business strategies such as digital marketing, inventory planning, and workforce scheduling. This study affirms that applying web scraping to customer reviews can serve as a practical approach for business decision-making and competitor analysis.*

Kata Kunci: *Web Scraping, Shopee, Google Collaboratory, Margin of Error, Visualisasi Data, Looker Studio, Data Ulasan.*

Abstrak: Penelitian ini membahas tentang pola penjualan produk dan data ulasan pengunjung pada platform *e-commerce* Shopee, dengan memanfaatkan teknik *web scraping* pada data ulasan pelanggan sebagai *proxy* tanggal transaksi. Tujuan penelitian meliputi eksplorasi pola penjualan produk, identifikasi lonjakan penjualan menurut waktu, dan penyajian visualisasi interaktif sebagai alat dukung keputusan. Data dikumpulkan melalui *web scraping* berbasis JavaScript yang dijalankan pada developer console untuk mengekstrak ulasan pelanggan. Setelah itu data diproses pada tahap pra pemrosesan data menggunakan alat google colab dan library python, untuk selanjutnya diunggah ke aplikasi Looker Studio sebagai bentuk pembuatan dashboard interaktif. Visualisasi yang dihasilkan lewat looker studio mendukung optimasi strategi bisnis seperti: *digital marketing*, dan perencanaan stok Studi ini menegaskan *web scraping* pada ulasan pelanggan dapat menjadi pendekatan praktis untuk

PENDAHULUAN

Di era digital saat ini, perdagangan elektronik (*e-commerce*) telah mengalami pertumbuhan eksponensial yang sangat tinggi di Indonesia. Platform Shopee menempati posisi pertama sebagai *e-commerce* dengan jumlah pengguna aktif dan volume transaksi bulanan terbanyak di Indonesia. Kemudahan berbelanja melalui Shopee membuat masyarakat dapat mengubah gaya hidup yang sedang tren (Ony, 2023).

Pertumbuhan pesat ini tidak hanya menunjukkan perubahan pola belanja konsumen dari *offline* ke *online*, tetapi juga mendorong berbagai pelaku bisnis untuk beradaptasi dengan strategi pemasaran dan bisnis berbasis data. Pemahaman mendalam terhadap pola penjualan produk dan *customer behaviour* menjadi sangat penting bagi para pelaku bisnis untuk menyusun strategi bisnis yang tepat sasaran, terutama ketika menghadapi periode tertentu seperti hari raya Idul Fitri, Natal, dan Tahun Baru, sehingga butuh analisis data yang tepat dan data yang mumpuni untuk menjadi pebisnis yang sukses.

Namun, di tengah besarnya potensi data yang tersedia, platform *e-commerce* seperti Shopee memiliki keterbatasan dalam memberikan data pola penjualan bulanan dan pola waktu *online* konsumen pada setiap sektor produk, yang lebih banyak disajikan kepada para pebisnis yang telah lama berbisnis di platform Shopee dalam bentuk *dashboard* yang hanya bisa diakses oleh pebisnis tersebut. Sehingga tidak adanya data publik yang mendukung analisis pola penjualan dan aktivitas konsumen yang akurat dan lengkap tiap produk.

Dengan adanya permasalahan di atas maka dibutuhkan teknik web scraping untuk mendapatkan data tanggal *review* konsumen pada website Shopee. Menurut *Data Preparation* untuk *Machine Learning* dan *Deep Learning* “*Web scraping* adalah teknik otomatisasi yang digunakan untuk mengambil atau mengekstrak data dari situs web. Dengan menggunakan *web scraping*, kita dapat mengumpulkan informasi dari halaman web yang ditampilkan kepada pengguna melalui *browser*.” (Rianto, 2024).

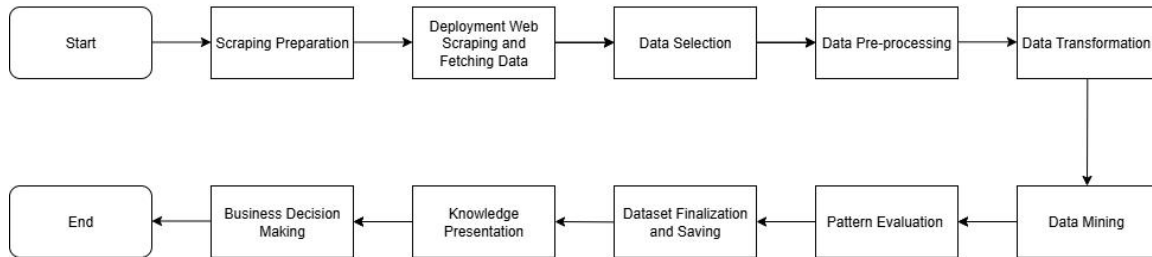
Dengan teknik *web scraping* pada data *review* pelanggan Shopee, teknik dapat menjadi solusi untuk mendapatkan data bagi para pebisnis, tanpa perlu memulai bisnisnya terlebih dahulu. Melalui teknik ini, data berupa tanggal publikasi *review* pelanggan, dapat dimanfaatkan sebagai *proxy* untuk memperkirakan pola aktivitas pengguna dan pola penjualan produk, sehingga diharapkan dapat memberikan gambaran yang lebih akurat mengenai *user behaviour* dari waktu ke waktu. Dalam penelitian ini, penulis menggunakan bahasa pemrograman Javascript untuk melakukan proses ekstraksi data (*web scraping*) dari halaman produk dan toko di platform *e-commerce* Shopee.

Diharapkan hasil dari penelitian ini tidak hanya memberikan kontribusi akademik melalui publikasi ilmiah, tetapi juga dapat diaplikasikan secara praktis oleh pelaku usaha mikro kecil menengah (UMKM), para pelaku bisnis, maupun para *digital marketer* yang mau memulai bisnis atau mengetahui *customer behaviour* yang lebih mendalam sebagai bagian dari strategi bisnis.

METODE

Penelitian menggunakan pendekatan kuantitatif berbasis data di bidang *e-commerce*, dengan fokus pada analisis data ulasan produk di Shopee untuk menentukan waktu optimal dalam strategi bisnis seperti digital marketing, *supply chain*, dan manajemen sumber daya manusia. Diagram alir penelitian ini merepresentasikan keseluruhan tahapan kerja dalam studi *data review*

produk di shopee berbasis *KDD* Berikut adalah diagram alir atau *flowchart* lengkap yang menjabarkan setiap langkah secara terperinci.



Gambar 1. Diagram Alir
(Sumber: Hasil Penelitian, [Draw.io](https://draw.io), 2025)

Diagram alir yang peneliti buat, disusun menggunakan *tools* dari *website* draw.io yang dirancang untuk menggambarkan proses secara sistematis dan mudah dipahami. Dalam penelitian ini peneliti memastikan kualitas dan relevansi data dalam penelitian, dengan menggunakan sejumlah kriteria seleksi yang bertujuan memperoleh data yang sesuai dengan tujuan analisis yang diinginkan. Kriteria-kriteria tersebut dirancang untuk menangkap pola dan volume penjualan sebagai data, serta memungkinkan analisis strategi bisnis di masa depan. Penjelasan lebih rinci mengenai masing-masing kriteria yang digunakan dalam pengambilan data dapat dilihat pada tabel berikut ini:

Tabel 1. Kriteria Pengambilan Data Penelitian

No	Kriteria	Penjelasan	Tujuan
1	Jumlah review ≥ 500	Produk yang dipilih harus memiliki minimal 500 <i>review</i> dan maksimal 5000 <i>review</i> dari halaman produk	Menjamin kuantitas data yang memadai
2	Riwayat penjualan aktif sejak november 2024	Toko penjual harus aktif melakukan transaksi sejak <i>quarter</i> ke 4, november 2024	Memastikan data yang dikumpulkan mencerminkan tren penjualan yang lengkap secara visual
3	Mendapatkan tag sebagai <i>brand official</i> “Shopee Mall”	<i>Brand</i> harus memiliki tag “Shopee Mall” sehingga <i>brand</i> sudah melalui verifikasi lebih lanjut oleh <i>team</i> shopee.	<i>Brand</i> telah dipastikan aman dari sisi penjualan, maupun hal-hal yang dikategorikan menyalahi aturan dan kebijakan toko di <i>platform</i> shopee, seperti “ <i>Bot buyers</i> ” dan lain sebagainya.

HASIL DAN PEMBAHASAN

Implementasi *Web Scraping* dan Ekstraksi *Dataset*

Implementasi proses *web scraping* dilakukan dengan proses ekstraksi *data review* dari platform Shopee, dengan terlebih dahulu dilakukan serangkaian langkah persiapan yang terstruktur guna memastikan bahwa proses *web scraping* dapat berjalan secara optimal, efisien, dan tidak terhambat oleh sistem keamanan yang diterapkan oleh *platform* tersebut. Persiapan ini mencakup pembuatan dan autentikasi akun Shopee, identifikasi target situs produk yang akan

diambil, hingga penyesuaian *tools* serta program yang akan digunakan dalam proses *scraping*.

Berikut adalah langkah langkah dalam implementasi web scraping sebagai berikut:

1. Persiapan akun dan akses situs shopee

Hal pertama yang dilakukan dalam proses melakukan web scraping ini adalah memiliki akses *user login* dengan membuka halaman utama Shopee melalui alamat “https://shopee.co.id/buyer/login” dan memilih opsi login yang diinginkan. Dengan memastikan bahwa sesi *login* akun masih berlaku dan semua kredensial terdaftar terdeteksi oleh sistem, risiko kegagalan otentikasi karena dianggap *bot*, maupun pemblokiran akses oleh mekanisme keamanan *platform* dapat di minimalisir atau bahkan ditiadakan.

2. Otentikasi dan validasi pengguna

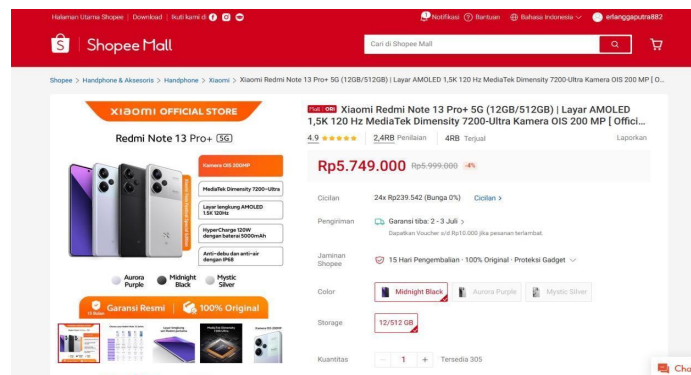
Setelah berhasil *login*, saat pertama kali mengakses halaman produk yang ingin di scraping, platform Shopee seringkali menampilkan validasi melalui verifikasi manusia (*human verification*) dengan metode *Puzzle* CAPTCHA. Pengguna wajib menyelesaikan *puzzle* CAPTCHA ini sekali, hal ini wajib dilakukan sebelum melakukan *web scraping* untuk menghindari kegagalan (status 403), karena dianggap sebagai *bot*. Sehingga setelah *puzzle* CAPTCHA terisi dan sesi dikonfirmasi *script scraping* dapat berjalan dengan baik.

3. Pemilihan *website* produk

Berikut adalah tabel daftar produk dan gambaran produk yang akan digunakan sebagai objek *scraping*:

Tabel 2. Kategori Produk

No	Kategori Produk	Brand	Produk
1	Smartphone	Xiaomi	Xiaomi Redmi Note 13 Pro+ 5G



Gambar 2. Tampilan Produk Xiaomi Redmi Note 13 Pro+
(Sumber: Xiaomi Official Store, Shopee, 2025)

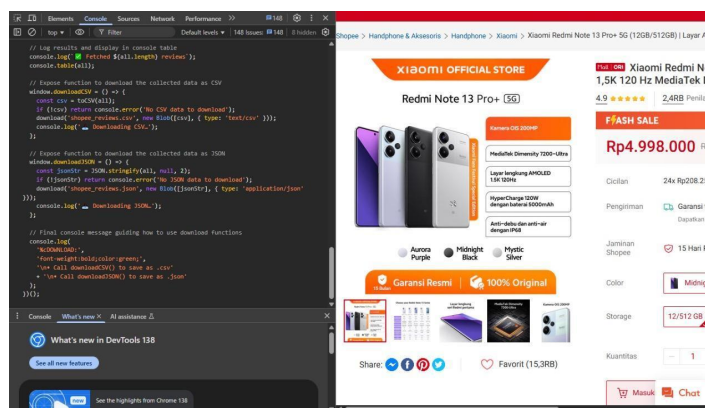
4. Implementasi *web scraping* dan ekstraksi *dataset*

Tahapan ini menjelaskan langkah-langkah teknis yang dilakukan untuk melakukan *scraping* data review produk dari halaman Shopee.

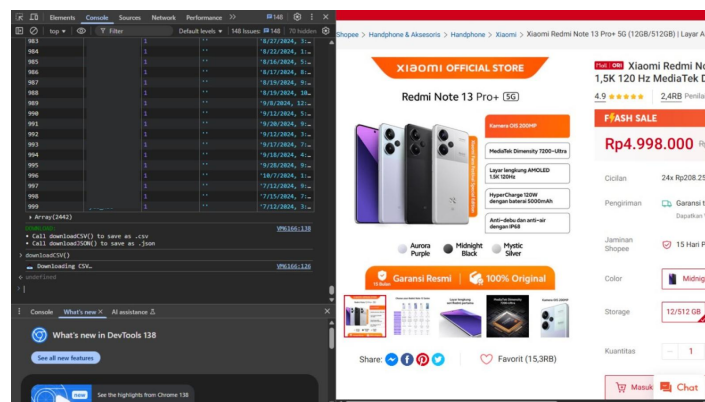
A. Proses *web scraping*

Proses dimulai dengan membuka halaman produk yang telah ditentukan. Beberapa produk yang menjadi objek *scraping* antara lain smartphone dari brand Xiaomi. Setelah halaman produk berhasil dibuka dan pengguna telah login ke akun Shopee menggunakan metode autentikasi, proses *scraping* langsung dapat dilakukan dengan masuk ke mode inspeksi *browser* (*Inspect Element*) dan berpindah ke *tab console*. Di dalam *tab console* tersebut, kode *scraping* disalin dan ditempelkan, kemudian dijalankan, lalu hasil file akan di simpan.

Berikut adalah gambaran langkap dari eksekusi *code* dan *deployment* pada produk Xiaomi Redmi Note 13 Pro+:



Gambar 3. *Code Deployment* Pada Halaman Produk
(Sumber: Xiaomi Official Store, Shopee, 2025)

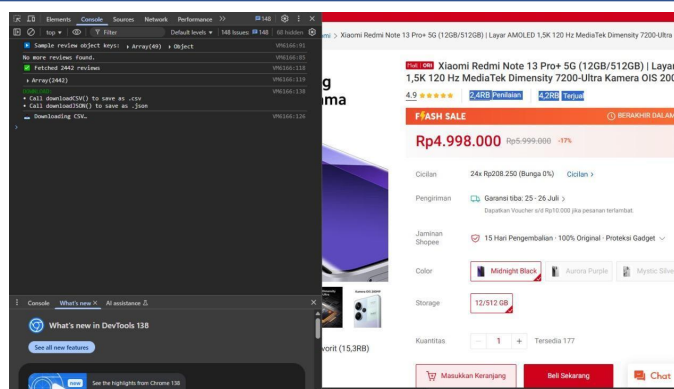


Gambar 4. Tampilan Hasil Scraping Pada Halaman Produk
(Sumber: Xiaomi Official Store, Shopee, 2025)

B. Proses data checking

Setelah proses web scraping berhasil dijalankan menggunakan *console*, langkah selanjutnya yang penting adalah melakukan verifikasi kesesuaian antara jumlah *review* yang berhasil diambil oleh skrip (*output scraping*) dan jumlah review yang ditampilkan secara publik pada halaman Shopee. Hal ini dilakukan untuk memastikan bahwa data yang diperoleh benar-benar merepresentasikan keseluruhan data yang tersedia secara aktual. Langkah pengecekan dilakukan dengan mencatat *total review* yang tampil di Shopee, yang tertera pada bagian "penilaian dan *review*" produk. Selanjutnya, total ini dibandingkan dengan jumlah *output* pada hasil ekstrasi *web scraping* yang terlihat pada di tabel *console*.

Berikut adalah gambaran langkap dari *Data Scraping Checking* pada produk Xiaomi Redmi Note 13 Pro+:



Gambar 5. Tampilan *Data Scraping Checking* Pada Halaman Produk
(Sumber: Xiaomi Official Store, Shopee, 2025)

C. Proses pengunduhan data

Tahapan selanjutnya data *review* yang berhasil dikumpulkan dan ditampilkan melalui *console browser* adalah menyimpan data tersebut ke dalam *file* yang dapat diolah lebih lanjut. *Script scraping* menyediakan dua metode untuk mengunduh hasil, yaitu “downloadCSV()” dan “downloadJSON()”. Untuk keperluan analisis dan visualisasi data, fungsi “downloadCSV()” digunakan karena hasilnya langsung dapat dibuka dengan *spreadsheet* atau diproses dalam bahasa pemrograman Python dan *library* nya, setelahnya *file* dapat dinamakan sesuai dengan nama produk atau nama lain yang diinginkan, sebagai contoh peneliti menggunakan nama *file* “Data Scraping - Xiaomi Redmi Note 13 Pro+ 5G 12GB-512GB.csv” sebagai nama *file* dari *review* produk yang di unduh. Proses ini memungkinkan pengguna untuk menyimpan data dengan rapi dan terorganisir, khususnya jika melakukan *scraping* terhadap beberapa produk dari kategori berbeda.

id	username	rating	comment	timestamp
1	y*****	1	Pesanan diterima lengkap, packing rapi dan aman, bubble wrap berapap. Produk bisa digunakan, update hyper OS. Pengiriman dan pengiriman cepat. Bekerja sesuai. Semoga semakin sukses. Kondisi barang dalam keadaan mulus dan lengkap. Fitur dan fungsi sesuai yang diiklankan. Keseluruhan memuaskan.	3/25/2024, 3:36:46 PM
2	g*****	1	Stok adalah bukannya Xiaomi Redmi Note 13 Pro+ 5G (12GB/512GB) dan seri ini adalah yang saya butuhkan saat ini untuk menunjang pekerjaan saya. Terima kasih.	11/11/2024, 22:11:08
3	g*****	1	Penyiriman pertama kali ke online, sudah dicoba 2 hari ini hp berfungsi dg baik selagi dalam masa garansi awal pengiriman cepat dan per. Fitur terbaik yang bisa dan mudah. Sesuai dengan harga yang sesuai.	3/17/2024, 8:55:53 AM
4	g*****	1	Jarum baru kali pertama beli hp secara online. Adminnya baik dan ramah cepat respon. Pengiriman cepat lagi. Ah pasti jadi langganan, sangat ar. Kondisi barang dalam keadaan mulus dan lengkap. Keseluruhan memuaskan.	5/13/2024, 8:52:03 AM
5	g*****	1	Pengiriman dan pengiriman cepat banget. Request dikirim sebelum jam 10an. Uda 2 kali beli disini dan semuanya oke. Keunggulan dibanding di kamarnya, tapi ini budayanya cuma karena yg lebih dan cepat buat kerja. Hp ini ude keren. Fitur Terbaik sesuai dengan yang diiklankan. Sesuai dengan harga yang sesuai.	7/8/2024, 11:11:29
6	g*****	1	Penyiriman cepat dan aman. Kondisi barang dalam keadaan mulus dan lengkap. Fitur dan fungsi sesuai yang diiklankan. Keseluruhan memuaskan.	5/11/2024, 11:55:14
7	g*****	1	Kontak sudah dibelikan buat main game. Fitur Terbaik yang bisa dan mudah. Sesuai dengan harga yang sesuai.	4/15/2024, 3:42:02 PM
8	g*****	1	Sampai cepat banget pengiriman dan barang aman nyampe di kondisi baik, hp canggih tp harga murah. Sesuai dengan harga yang sesuai.	5/2/2024, 18:31:22
9	g*****	1	Mudah-mudahan awal. Pengiriman cepat kaya kilat.	3/19/2024, 11:58:50

Gambar 6. Tampilan CSV *Data Scraping* Xiaomi Redmi Note 13 Pro+
(Sumber: Xiaomi Official Store, Shopee, 2025)

Implementasi Tahapan *Knowledge Discovery in Databases (KDD)*

Implementasi *Data Selection*

Tahapan pertama dalam proses *Knowledge Discovery in Database (KDD)* adalah *data selection*, tahap *data selection* dilakukan untuk memastikan bahwa data yang digunakan dalam penelitian benar-benar relevan dengan periode analisis yang ditetapkan, yaitu tahun 2024 hingga 2025. Data yang digunakan oleh peneliti dalam laporan ini adalah data file *comma-separated values (CSV)* dari tahap *web scraping* pada produk Xiaomi Redmi Note 13 Pro+ 5G 12GB-512GB.

Implementasi tahap ini dilakukan menggunakan google colab dengan bantuan *library* Pandas, yang berfungsi untuk mengelola dan memproses data dalam format *comma-separated values* (CSV) secara efisien. Pemanfaatan google colab dipilih karena menyediakan *environment* komputasi berbasis *cloud* yang mendukung analisis data tanpa perlu instalasi perangkat lunak tambahan, sedangkan *library* Pandas memungkinkan peneliti melakukan manipulasi data secara fleksibel, mulai dari pembacaan *file* hingga proses penyaringan data berdasarkan kriteria tertentu. Proses implementasi tahap *data selection* ini dilakukan melalui serangkaian langkah-langkah berikut:

1. Proses *writing code* pada *notebook* google colab
2. Proses mengunggah *file* menggunakan *function* “files.upload()”
3. Proses eksekusi *code* menggunakan *library* pandas dan io
4. Proses menampilkan *preview* data menggunakan *function* “display()”
5. Proses pengunduhan data menggunakan *function* "files.download()"

Implementasi *Preprocessing* (Data Cleaning)

Tahapan *Preprocessing* (Data Cleaning) merupakan proses penting dalam metode *Knowledge Discovery in Databases* (KDD) yang bertujuan untuk menyiapkan data agar siap digunakan pada tahap analisis berikutnya. Terdapat penggunaan *data cleaning* dan *data reduction* dalam implementasi tahap *Preprocessing* ini. Sejumlah proses *preprocessing* data diterapkan untuk meningkatkan tingkat akurasi model prediksi. (Gori, et, al. 2024) Pada penelitian ini, tahap *Preprocessing* dilakukan untuk memastikan bahwa data hasil *web scraping* dari platform Shopee telah bersih dan relevan.

Proses ini dilakukan menggunakan google colab melalui beberapa langkah sistematis sebagai berikut:

1. Proses *writing code* pada *notebook* google colab
2. Proses mengunggah *file* menggunakan *function* “files.upload()”
3. Proses eksekusi *code* menggunakan *library* pandas dan io
4. Proses menampilkan *preview* data menggunakan *function* “display()”
5. Proses pengunduhan data menggunakan *function* "files.download()"

Implementasi *Data Transformation*

Tahap *data transformation* bertujuan untuk mengubah format data hasil *preprocessing* menjadi bentuk yang bisa dinilai, yaitu dalam *format volume* ulasan per bulan. Pada penelitian ini, proses transformasi dilakukan dengan mengubah data *timestamp* menjadi *volume* bulan dan tahun dalam bentuk *alphabetical month name* yang menghitung kemunculan jumlah bulan dan tahun pada setiap *entries* di kolom *timestamp* tersebut.

Proses ini dilakukan menggunakan Google Colab melalui langkah-langkah berikut:

1. Proses *writing code* pada *notebook* google colab
2. Proses mengunggah *file* menggunakan *function* “files.upload()”
3. Proses eksekusi *code* menggunakan *library* pandas dan io
4. Proses menampilkan *preview* data menggunakan *function* “display()”
5. Proses pengunduhan data menggunakan *function* "files.download()"

Hasil *data transformation* ini menghasilkan *dataset* akhir yang telah diolah dalam bentuk kolom “date_timestamp” dan “date_volume”, yang memuat jumlah ulasan produk Xiaomi Redmi Note 13 Pro+ 5G 12GB-512GB berdasarkan periode waktu bulanan. Proses ini memastikan bahwa data mentah yang sebelumnya tidak terstruktur kini telah dikonversi menjadi format yang

lebih mudah dianalisis. Dengan demikian, hasil dari tahap *data transformation* ini menjadi landasan penting untuk mendukung tahap analisis dan pemodelan pada tahap selanjutnya dalam proses *Knowledge Discovery in Database (KDD)*.

Implementasi *Data Mining*

Data mining merupakan proses sistematis untuk menggali informasi berharga dari kumpulan data besar melalui tahapan terstruktur seperti yang dilakukan pada dua tahap *data mining* ini. Pada tahap pertama *feature enrichment*, data hasil sebelumnya disiapkan untuk membangun fondasi analisis yang sesuai dengan *library* dan menambahkan variabel penting seperti “base_proportion” untuk dasar proporsi dalam menghitung angka estimasi penjualan dari data *review*, dan “trend_factor” untuk menghitung perubahan *trend* pada *volume data*, sehingga pola perhitungan estimasi penjualan dari data *review* dapat diperkaya dengan konteks menggunakan variabel “weight” untuk menghitung nilai pengaruh yang digunakan untuk menentukan “allocated_extra” alokasi tambahan pada variabel data *estimated sales* dengan penambahan tepat sesuai bobot nilai setiap bulan, pada variabel “estimated_sales” dan menggambarkan estimasi penjualan yang sebenarnya pada setiap bulan. Pada tahap terakhir *margin of error (MOE)* menilai tingkat keandalan data. Proses ini dilakukan menggunakan google colab melalui langkah-langkah berikut:

1. Proses *writing code* pada *notebook* google colab
2. Proses mengunggah *file* menggunakan *function* “files.upload()”
3. Proses eksekusi *code* menggunakan *function* variabel *base proportion*, dan *factor trend*
4. Proses eksekusi *code* menggunakan *function* variabel *margin of error (MOE)*
5. Proses menampilkan *preview* data menggunakan *function* “display()”
6. Proses pengunduhan data menggunakan *function* “files.download()”

Implementasi *Pattern Evaluation*

Data mining merupakan proses sistematis untuk menggali informasi berharga dari kumpulan data besar melalui tahapan terstruktur seperti yang dilakukan pada dua tahap *data mining* ini. Pada tahap pertama *feature enrichment*, data hasil sebelumnya disiapkan untuk membangun fondasi analisis yang sesuai dengan *library* dan menambahkan variabel penting seperti “base_proportion” untuk dasar proporsi dalam menghitung angka estimasi penjualan dari data *review*, dan “trend_factor” untuk menghitung perubahan *trend* pada *volume data*, sehingga pola perhitungan estimasi penjualan dari data *review* dapat diperkaya dengan konteks menggunakan variabel “weight” untuk menghitung nilai pengaruh yang digunakan untuk menentukan “allocated_extra” alokasi tambahan pada variabel data *estimated sales* dengan penambahan tepat sesuai bobot nilai setiap bulan, pada variabel “estimated_sales” dan menggambarkan estimasi penjualan yang sebenarnya pada setiap bulan. Pada tahap terakhir *margin of error (MOE)* menilai tingkat keandalan data. Proses ini dilakukan menggunakan google colab melalui langkah-langkah berikut:

1. Proses *writing code* pada *notebook* google colab
 Pada tahap *pattern evaluation* penulisan kode dilakukan di google colab dengan beberapa fungsi dan helper function yang disusun untuk membaca variasi format tanggal, dan nilai MOE, untuk kemudian memvisualisasikan nya dalam tabel chart line, hal pertama adalah files.upload() menampilkan widget unggah berkas di google colab, pd.read_csv() membaca CSV ke DataFrame, robust_parse_date() untuk fungsi yang mencoba pd.to_datetime() dan

beberapa format eksplisit serta `re.search()` (regex) sebagai cadangan untuk mengonversi string tanggal menjadi objek `datetime`, `DataFrame.apply()` menerapkan parser per baris, `parse_moe_value()` dan string MOE menjadi nilai numerik dalam persen, `pd.to_numeric()` dan `np.nan` menangani konversi numerik dan nilai hilang. mengagregasi data per bulan menghasilkan `total_moe`, `mean_moe`, dan `count_rows`, sedangkan `plt.bar()` dan `plt.plot()` (matplotlib) menghasilkan visualisasi line chart total MOE dan garis mean MOE.

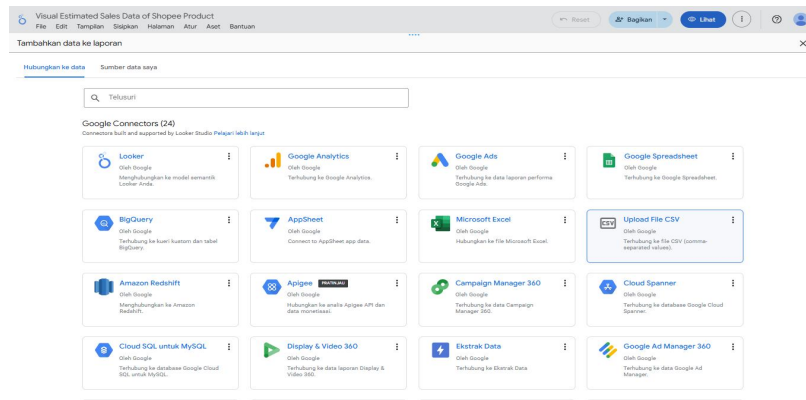
2. Proses eksekusi *code* menggunakan *library* pandas, matplotlib, IPython dan IO

Proses eksekusi kode dimulai melalui fungsi `upload_or_fallback()` yang bertugas menangani input data, di mana pengguna diminta mengunggah file CSV pada google colab. Data yang terbaca dikembalikan dalam bentuk `DataFrame` `df` beserta nama sumber `source_name`. Selanjutnya, sistem mendeteksi kolom MOE (`moe_col`) dan kolom waktu (`ts_col`) melalui pemeriksaan nama kolom secara bertingkat agar mampu menyesuaikan variasi penamaan. Fungsi `robust_parse_date(s)` digunakan untuk melakukan parsing tanggal secara berlapis, mulai dari penggunaan `pd.to_datetime()`, berbagai format tanggal eksplisit, hingga dua pola regex fallback untuk menangani variasi format seperti 'June 2025' atau '06/2025'. Hasil *parsing* disimpan pada kolom `ts_parsed`, sedangkan fungsi `parse_moe_value(x)` berfungsi menormalkan nilai MOE dengan menghapus tanda persen, menyesuaikan format desimal, dan mengkonversinya menjadi persen jika diperlukan, kemudian disimpan dalam kolom `moe_value`. Setelah itu, sistem menandai baris bermasalah di mana nilai MOE tersedia namun tanggal tidak terparse untuk membantu proses diagnosis. Tahap selanjutnya adalah agregasi bulanan yang menghasilkan `total_moe`, `mean_moe`, dan jumlah entri per bulan melalui metode `groupby()`. Hasil agregasi divisualisasikan menggunakan grafik kombinasi batang dan garis (`plt.bar` dan `plt.plot`) untuk menampilkan total serta rata-rata MOE per bulan. Pengaturan label, rotasi teks, legenda, dan grid diterapkan agar grafik tampak informatif dan terstruktur, sehingga mampu merepresentasikan pola dan tren MOE secara jelas berdasarkan periode waktu dari sumber data CSV.

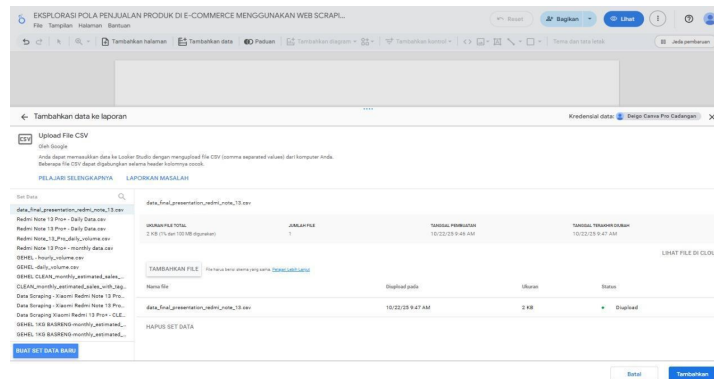
Implementasi *Knowledge Presentation*

Pada tahapan ini, dilakukan penyajian data dalam bentuk visualisasi lanjutan untuk mendukung pemahaman yang lebih mendalam terhadap informasi yang telah dikumpulkan. Tahap implementasi *knowledge presentation* dilakukan menggunakan google looker studio sebagai media visualisasi interaktif untuk menyajikan hasil analisis data. *Dashboard* yang dibuat menampilkan pola estimasi penjualan bulanan dalam bentuk grafik batang bar chart, yang menggambarkan fluktuasi volume penjualan produk Xiaomi Redmi Note 13 Pro+ 12/256 GB 5G dari Maret 2024 hingga Juni 2025. Visualisasi ini memudahkan analisis bisnis dalam memahami tren penjualan secara cepat dan intuitif untuk meneliti penjualan kompetitor maupun mempelajari produk yang akan dijual, di mana setiap *bar* pada *chart* merepresentasikan jumlah estimasi penjualan per bulan. Melalui pendekatan ini, informasi yang dihasilkan dari proses data mining dapat disajikan secara komprehensif dan mudah diinterpretasikan, sehingga mendukung proses pengambilan keputusan berbasis data dengan tampilan yang dinamis dan informatif.

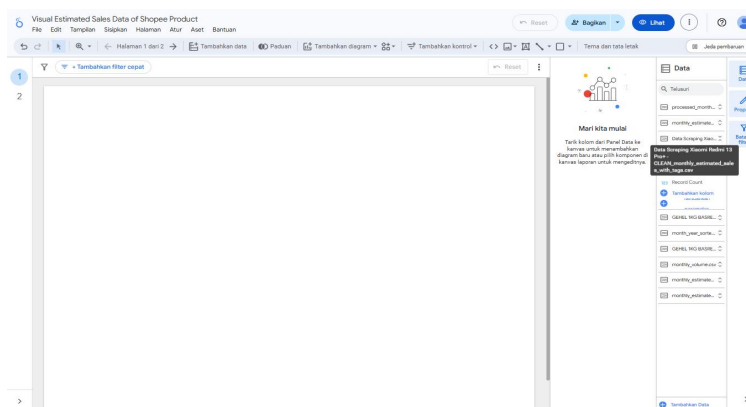
Berikut adalah gambaran dari penggunaan software looker studio untuk menghasilkan *knowledge presentation* dalam *bar chart*:



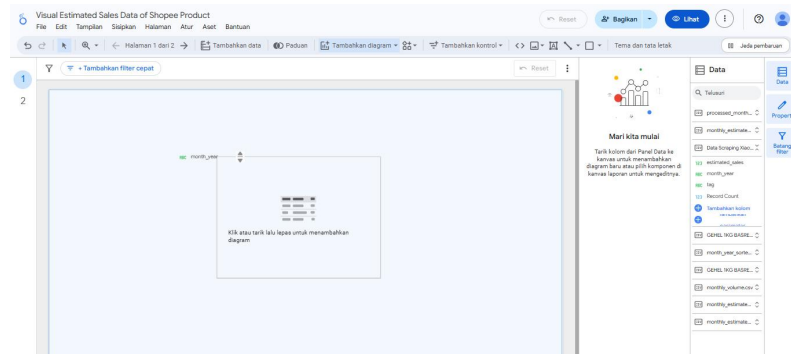
Gambar 7. Looker Studio *Upload Monthly Data*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



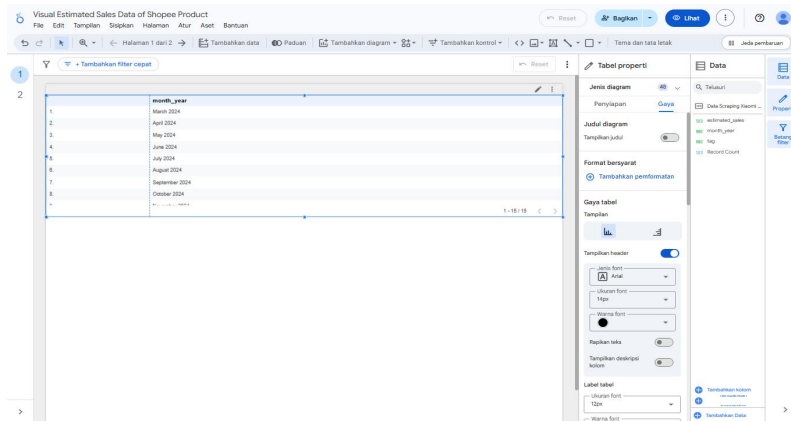
Gambar 8. Looker Studio *Load Monthly Data*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



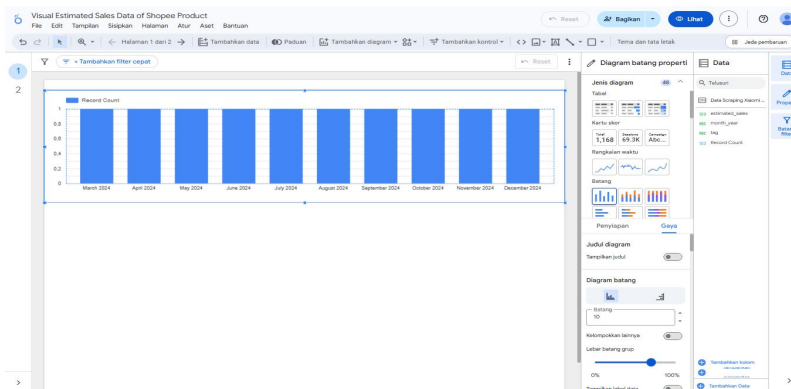
Gambar 9. Looker Studio *Blank Dashboard*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



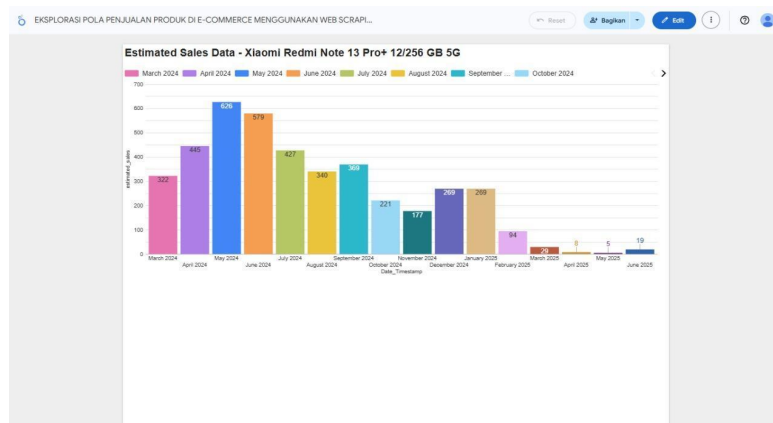
Gambar 10. Looker Studio *Drag Data*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



Gambar 11. Looker Studio *Monthly Data Table*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



Gambar 12. Looker Studio *Monthly Pra Bar Chart Visual*
(Sumber: Hasil Penelitian, Looker Studio, 2025)



Gambar 13. Looker Studio *Estimated Sales Bar Chart Visual*
(Sumber: Hasil Penelitian, Looker Studio, 2025)

Hasil Implementasi Strategi Bisnis Pada Platform Shopee.

Hasil Implementasi Perencanaan *Digital Marketing*

Menurut *Digital Marketing: Brand Images* mendefinisikan “*digital marketing* memainkan peran penting dalam membangun aktivasi merek. Karena, saat ini kita telah memasuki era digital, di mana semua orang sudah akrab dengan keberadaan internet, yang telah menjadi perilaku di masyarakat” (Sawhani, 2021).

Perencanaan *digital marketing* yang efektif membutuhkan dasar data yang kuat. Dalam konteks penelitian ini, estimasi pola penjualan setiap bulan pada periode yang ditentukan memiliki peran penting dalam menentukan waktu tayang iklan, penentuan anggaran, dan strategi promosi.

1. Penentuan Waktu Iklan Berdasarkan *Data Visual*

Dari data yang telah diberikan, pelaku usaha dapat melihat bulan dengan volume penjualan atau aktivitas tertinggi. Menggunakan *visual data*, kita dapat memilih dan menjadwalkan iklan produk maupun Shopee *live* yang disesuaikan dengan puncak aktivitas pengguna.

2. Pengaturan Anggaran Iklan

Dengan menganalisis hasil visual data, pebisnis dapat menentukan kapan iklan perlu ditingkatkan atau dikurangi. Misalnya pada bulan-bulan dengan penjualan tinggi, iklan ditayangkan secara masif dan diberi anggaran tinggi untuk memperkuat *brand awareness*. Alokasi penganggaran iklan yang optimal di Instagram tidak hanya berpengaruh pada kualitas visual dan banyaknya iklan yang ditampilkan tetapi juga berdampak pada persepsi nilai merek di mata konsumen (Arnoldi, 2025).

3. Penentuan Periode *Flash Sale*

Data penjualan bulanan dapat membantu dalam merencanakan periode *Flash Sale*. Misalnya, jika bulan tertentu menunjukkan tren estimasi penjualan tertinggi, maka pelaksanaan *Flash Sale* sebaiknya dilakukan pada bulan tersebut. Dalam *dashboard* Shopee, fitur “*Flash Sale Toko Saya*” memungkinkan penjual memilih sesi, menambahkan produk, dan menampilkannya secara strategis. Ini dapat dimanfaatkan untuk mendongkrak penjualan produk baru.

Berikut adalah gambaran dari halaman iklan dan halaman pengaturan *flash sale* dari shopee:

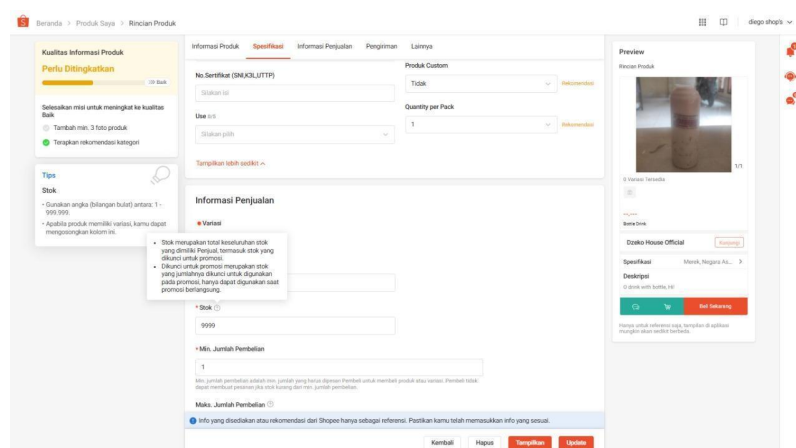
Gambar 14. Halaman *Dashboard* Iklan Produk Shopee
(Sumber: Hasil Penelitian, Shopee, 2025)

Gambar 15. Halaman *Dashboard* Iklan Shopee Live
(Sumber: Hasil Penelitian, Shopee, 2025)

Implementasi Perencanaan *Supply Chain*

Perencanaan *supply chain* yang efektif dimulai dari pemahaman terhadap perilaku permintaan produk yang dapat diperoleh melalui *dataset* hasil *scraping* seperti volume penjualan setiap bulan. Dengan menganalisis tren penjualan ini, pebisnis baru dapat memprediksi kebutuhan stok yang akurat dan menyesuaikannya melalui fitur pengaturan stok pada *dashboard* Shopee.

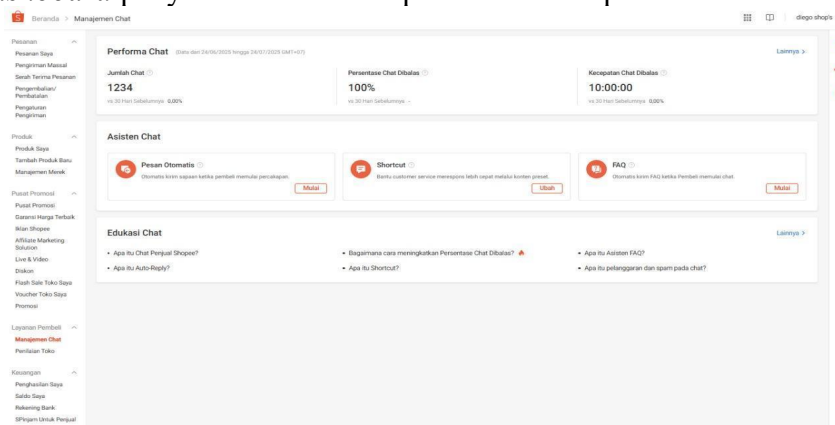
Berikut adalah gambaran dari halaman *dashboard* penyesuaian stok dan produk dari shopee:



Gambar 16. Halaman *Dashboard* Iklan Produk Shopee
(Sumber: Hasil Penelitian, Shopee, 2025)

Implementasi Perencanaan *Human Resources*

Dataset yang telah di-scraping dapat memberikan wawasan penting bagi pebisnis baru untuk merancang strategi *human resources* yang efisien dan berbasis data. Hasil pola data memungkinkan pebisnis mengetahui bulan-bulan dengan potensi interaksi tinggi dari pelanggan. Dari grafik, terlihat bahwa puncak penjualan dan aktivitas pengunjung terjadi di waktu tertentu, dengan demikian admin perlu dijadwalkan aktif pada periode tersebut. Berikut adalah gambaran dari halaman *dashboard* penyesuaian stok dan produk dari shopee:



Gambar 17. Halaman *Dashboard* Chat Shopee
(Sumber: Hasil Penelitian, Shopee, 2025)

KESIMPULAN

Berdasarkan hasil penelitian mengenai eksplorasi data *review* untuk mengidentifikasi pola penjualan produk di platform Shopee, dapat disimpulkan bahwa data ulasan pelanggan dapat digunakan sebagai indikator dalam mengestimasi pola penjualan di *e-commerce*. Melalui teknik *web scraping* terhadap estimasi pola penjualan setiap bulan pada periode yang ditentukan, penelitian ini berhasil memperkirakan estimasi pola penjualan bulanan suatu produk tanpa perlu memiliki data nya secara internal, dan dapat mengambil data secara eksternal. Visualisasi estimasi yang kemudian divisualisasikan dalam Looker Studio, dilengkapi dengan grafik bar chart memungkinkan pelaku usaha untuk tidak hanya mengenali pola penjualan tapi juga merancang strategi pemasaran dan manajemen stok serta perencanaan strategi bisnis secara

online lainnya.

Adapun *Dataset* yang dihasilkan dapat memberikan landasan untuk pengambilan keputusan bisnis, seperti penjadwalan kampanye iklan, pengadaan stok menjelang lonjakan permintaan, serta penyesuaian jam operasional toko berdasarkan aktivitas pelanggan, tanpa adanya bias persepsi. Dengan demikian, penelitian ini menempatkan eksplorasi data *review* melalui *web scraping* untuk dapat menjadi solusi alternatif bagi usaha mikro kecil menengah (UMKM) dan pelaku bisnis yang memiliki keterbatasan akses terhadap data penjualan pada e-commerce dan mendukung strategi bisnis berbasis data.

REFERENSI

- Alfiah, F., Agung, Y. N., Ade S., Indo I., Wiranti S. U., Hamzah N. F., Norbertus, T. S. S., Tugiman, Mohammad, B., Max, T. A. C. W., Erna, A., Patria, A., Siti M., Nasril, S., Muhamad, Y., Inda R. S., & Aqil, F. (2025). *Pemodelan dan Visualisasi Data*. Banten: Sada Kurnia Pustaka.
- Alghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *JURNAL ILMIAH INFORMATIKA*, 9(02), 75–81. <https://doi.org/10.33884/jif.v9i02.3755>
- Aritonang, M. A. S., Imanuel, Z., Castaka, A. S., Isa, A., Rudy, C., Fendi, H., Dadan, S. R., Valian, Y. P. A., Triawan, A. C., Arif, M. W., Retno, D., Yesri, E., Indrawan, A. S., & Moh, S. (2024). *Dasar Pemrograman Python*. Sumatra Barat: Yayasan Tri Edukasi Ilmiah.
- Arnoldi, G. S. (2025). Pengaruh alokasi anggaran iklan digital terhadap persepsi nilai merek pada konsumen generasi Z: Studi kuantitatif di media sosial Instagram. *Musytari: Jurnal Manajemen, Akuntansi, dan Ekonomi*, 20(2), 51–60. <https://doi.org/10.2324/waajnv75>
- Azora, P. (2021). Analisis quick count dengan menggunakan metode stratified random sampling studi kasus Pemilu Gubernur Kalimantan Barat 2018. *BIMASTER: Buletin Ilmiah Matematika, Statistika dan Terapannya*, 10(1), 43–50. <https://doi.org/10.26418/bbimst.v10i1.44666>
- Baskara, R. (2022). Implementasi Web Scraping pada Media Sosial Instagram. *Jurnal Informatika UII AUTOMATA*, 3(1). <https://journal.uii.ac.id/AUTOMATA/article/view/21918/12108>
- Djiwadikusumah, F., Irawan, G. H., & al-Fadilah, R. H. (2021). Web Scraping Situs E-commerce Menggunakan Teknik Parsing DOM. *Jurnal Siliwangi Seri Sains dan Teknologi*, 7(2), 52–57. <https://doi.org/10.37058/jssainstek.v7i2.4223>
- Febryana, S., Wulandari, I. A., & Prabowo, D. (2024). Rancang bangun aplikasi pengolahan data inventory barang berbasis desktop pada CV. Hamim Group. *Jurnal Mahasiswa Sistem Informasi (JMSI)*, 6(1), 24-35. <https://doi.org/10.24127/jmsi.v6i1.7553>
- Gehel. (2025). Gehel Basreng Cikruh 1 KG. Shopee. Retrieved 2025, from [https://shopee.co.id/GEHEL-1KG-\(BASRENG-CIKRUH-DAN-BASRENG-CIAMIK\)-Camilan-Makanan-Ringan-i.120552964.20981025148](https://shopee.co.id/GEHEL-1KG-(BASRENG-CIKRUH-DAN-BASRENG-CIAMIK)-Camilan-Makanan-Ringan-i.120552964.20981025148)
- Indrayani N. L. A. (2022). Penerapan sistem enterprise resource planning (ERP) pada perusahaan jasa konstruksi. *CRANE: Civil Engineering Research Journal*, 3(2), 11–16. <https://doi.org/10.34010/crane.v3i2.8159>
- Kusumo, S., & Somya, R. (2022). Penerapan Web Scraping Deskripsi Produk Menggunakan Selenium Python Dan Framework Laravel. *Jurnal Teknik Informatika dan Sistem Informasi*, 9(4), 3426–3435. <https://doi.org/10.35957/jatisi.v9i4.2727>
- Maulana, Y. (2024). *Belajar Ngoding Python Untuk Data Science*. Mobidu.

-
- Najamuddin, J., Satoto, B. D., & Negara, Y. D. P. (n.d.). Implementasi teknik web scraping untuk pengumpulan data laporan keuangan perusahaan di Bursa Efek Indonesia (IDX). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2), 2789–2795. <https://doi.org/10.36040/jati.v9i2.13070>
- Ony, D. S. S., & Agus, S. (2023). Pengaruh Shopee PayLater, Promosi Penjualan, dan Gaya Hidup Hedonisme Terhadap Perilaku Konsumtif Mahasiswa Penulis. *Journal Manajemen dan Bisnis (Univ. Teknologi Sumbawa)*, Vol.6 No.2, 2023. DOI:10.37673/jmb.v6i2.3584
- Rianto, & Paulus Insap Santosa. (2025). *Data Preparation untuk Machine Learning & Deep Learning*. Yogyakarta: ANDI.
- Rizquina, A. Z., & Ratnasari, C. I. (2023). Implementasi Web Scraping untuk Pengambilan Data Pada Website E-Commerce. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 5(4), 377–383. <https://doi.org/10.47233/jteksis.v5i4.913>
- Sawhani, K. D. (2021). *Digital Marketing: Brand Images*. Surabaya: SCOPINDO MEDIA PUSTAKA.
- Setia, W., & Muhammad, F. (2023). Web Scraping Untuk Pengambilan Data Transaksi Penjualan Pada Markeplace Toko Online Jogja batik. *JIIFKOM (Jurnal Ilmiah Informatika dan Komputer)*, 2(2):14-17. DOI:10.51901/jiifkom.v2i2.254
- Sopian, A., Dharmalau, A., & Alpindo. (2022). Pemanfaatan Teknik Web Scraping Python untuk Sistem Pencarian Produk di Toko Online. *Jurnal Elektro & Informatika Swadharma (JEIS)*, 2(2). <https://doi.org/10.56486/jeis.vol2no2.169>
- Suryadi, A., W., S. A., Alfa'inna, N., & E., H. H. (2023). Implementasi Web Scraping dan Sentiment Analysis terhadap Berita Menggunakan Machine Learning. *Jurnal Swabumi*, 11(1), 28~34. <https://doi.org/10.31294/swabumi.v11i1.15145>
- Swastika R., Siti M., Ferry S., Muhamad, M., & Sri, I. (2023). *IMPLEMENTASI DATA MINING (Clustering, Association, Prediction, Estimation, Classification)*. Indramayu: Adab.
- Ulfah, A., & Najiah, I. (2023). Implementasi Web Scraping Pada Situs Jurnal SINTA Menggunakan Framework Selenium WebDriver Python. *JIKA (Jurnal of Informatics)*, 7(1), 29–36. <http://dx.doi.org/10.31000/jika.v7i1.7037>
- Wahyudi, R., & Kusumawardhana, G. (n.d.). Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine. *Jurnal Informatika*, 8(2), 200–207. <https://doi.org/10.31294/ji.v8i2.9681>
- Wardhana, R. G., Wang, G., & Sibuea, F. (2023). penerapan machine learning dalam prediksi tingkat kasus penyakit di Indonesia. *Journal of Information System Management (JOISM)*, 5(1), 40-45. <https://doi.org/10.24076/joism.2023v5i1.1136>
- Xiaomi Redmi Note 13 Pro+ 5G (12GB/512GB) | Layar AMOLED 1,5K 120 Hz MediaTek Dimensity 7200-Ultra Kamera OIS 200 MP [Official Store]. (2025). Shopee. Retrieved 2025, from [https://shopee.co.id/Xiaomi-Redmi-Note-13-Pro-5G-\(12GB-512GB\)-Layar-AMOLED-1-5K-120-Hz-MediaTek-Dimensity-7200-Ultra-Kamera-OIS-200-MP-Official-Store--i.51925611.25366617978](https://shopee.co.id/Xiaomi-Redmi-Note-13-Pro-5G-(12GB-512GB)-Layar-AMOLED-1-5K-120-Hz-MediaTek-Dimensity-7200-Ultra-Kamera-OIS-200-MP-Official-Store--i.51925611.25366617978)
- Yuda, M. A. D. (2025). Implementasi Web Scraping Untuk Ekstraksi Data Penjual Dan Produk Panel Surya Di E-Commerce. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 5(1). <https://doi.org/10.24002/konstelasi.v5i1.11751>