

Perancangan Data Pipeline Untuk Analisis Pola Makan Sehat Sebagai Upaya Pencegahan Gizi Buruk di Indonesia

Ghaida Putri Yunisa¹, Fitri Akmaliah², Allya Putri Nadila Agustin³, Yusma Cantika Parhati⁴, Chaerur Rozikin⁵

^{1,2,3,4,5} Universitas Singaperbangsa Karawang, Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang 41361, Indonesia

e-mail: 2210631170022@student.unsika.ac.id¹, 2210631170020@student.unsika.ac.id², 2210631170111@student.unsika.ac.id³,

2110631170117@student.unsika.ac.id⁴, chaerur.rozikin@staff.unsika.ac.id

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi : 15 Oktober 2025

Revisi Akhir : 04 November 2025

Diterbitkan Online : 10 November 2025

Kata Kunci:

Data Pipeline, Pola Makan Sehat, Gizi Buruk, Nutrisi Harian, Kaggle

Korespondensi:

Telepon / Hp : +6281211876248

E-mail :

2210631170022@student.unsika.ac.id

ABSTRAK

Masalah gizi di Indonesia, termasuk *stunting* dan obesitas sentral, masih menjadi tantangan kesehatan yang signifikan akibat pola makan yang tidak seimbang dan rendahnya pengetahuan masyarakat tentang nutrisi. Penelitian ini bertujuan merancang data *pipeline* untuk mengintegrasikan, membersihkan, dan menyiapkan data gizi dari *dataset* publik, seperti *Kaggle*, sehingga analisis pola makan sehat dapat dilakukan secara komprehensif. *Pipeline* yang dikembangkan mencakup tahap data *ingestion*, data *cleaning* dan *preprocessing*, transformasi data, penyimpanan, serta analisis dan visualisasi data. Hasil implementasi menunjukkan bahwa setiap individu memiliki variasi asupan nutrisi harian yang berbeda, yang kemudian dibandingkan dengan standar kecukupan gizi harian WHO/Kemkes. Sistem memberikan ringkasan status gizi individu dalam kategori “Seimbang”, “Perlu Perbaikan”, dan “Kurang Sehat”, dengan mayoritas responden masuk kategori “Perlu Perbaikan” (59,6%). Temuan ini menekankan perlunya intervensi berbasis data untuk meningkatkan pola makan masyarakat. *Dataset* yang dihasilkan bersih, terstruktur, dan siap digunakan untuk analisis lebih lanjut, sehingga dapat mendukung tenaga kesehatan, pembuat kebijakan, dan peneliti dalam merancang program pencegahan gizi buruk yang tepat sasaran.

1. PENDAHULUAN

Masalah gizi di Indonesia masih menjadi tantangan besar yang memengaruhi kualitas kesehatan dan produktivitas masyarakat. Menurut Survei Status Gizi Indonesia (SSGI) 2024, prevalensi *stunting* nasional mencapai 19,8% pada tahun menurun dari 21,5% pada 2023. Meskipun terjadi penurunan, angka ini masih jauh dari target jangka panjang pemerintah untuk menurunkan *stunting* menjadi 14,2% pada 2029 [1]. Selain masalah kekurangan gizi, fenomena obesitas sentral juga menjadi tantangan kesehatan publik. Data Survei Kesehatan Indonesia (SKI) 2023 mencatat prevalensi obesitas sentral pada penduduk usia 15 tahun ke atas mencapai 36,8%, meningkat dari 31% pada 2018 [2]. Konsumsi sayur dan buah per kapita di Indonesia pada 2023 hanya mencapai 240,5 gram per hari, jauh di bawah rekomendasi gizi seimbang [3].

Faktor penyebab utama dari masalah gizi antara lain pola makan yang tidak seimbang, rendahnya pengetahuan masyarakat tentang kandungan gizi makanan, serta tingginya konsumsi makanan cepat saji. Gizi seimbang adalah pola konsumsi makanan yang mencukupi kebutuhan tubuh akan protein, karbohidrat, vitamin, mineral, dan nutrisi lainnya secara proporsional, sesuai dengan aktivitas fisik sehari-hari. Selain itu, kebersihan makanan juga perlu diperhatikan. Asupan nutrisi yang seimbang sangat penting untuk menjaga kesehatan tubuh dan mendukung kondisi mental yang optimal [4].

Meskipun data terkait gizi dan konsumsi masyarakat tersedia melalui berbagai sumber, seperti SSGI, SKI, data konsumsi per kapita dari Badan Pusat Statistik (BPS), serta data produksi pangan dari Kementerian Pertanian, data ini masih tersebar dan belum terintegrasi. Fragmentasi data ini menjadi kendala dalam analisis pola makan masyarakat secara komprehensif, sehingga identifikasi kelompok berisiko gizi buruk, perumusan kebijakan berbasis bukti, dan penentuan prioritas intervensi gizi menjadi kurang akurat dan kurang optimal.

Berdasarkan kondisi tersebut, perancangan data *pipeline* menjadi strategi penting untuk mengelola dan menganalisis informasi gizi secara sistematis. Dengan *pipeline* yang efektif, data dari satu sumber publik, seperti *dataset Kaggle*, dapat dibersihkan, disusun, dan diolah sehingga analisis pola makan masyarakat dapat dilakukan secara menyeluruh, mulai dari identifikasi asupan nutrisi individu hingga penentuan risiko gizi buruk pada kelompok tertentu. Implementasi *pipeline* yang baik tidak hanya mempermudah analisis statistik dan visualisasi data, tetapi juga dapat menyediakan informasi yang dapat digunakan oleh tenaga kesehatan, pembuat kebijakan, maupun peneliti untuk merancang program intervensi gizi yang lebih tepat sasaran dan berbasis data [5].

Penelitian ini bertujuan untuk merancang data *pipeline* yang efektif dalam membersihkan, menyiapkan, dan mengolah data pola makan dari *dataset* publik seperti *Kaggle*. Hasil penelitian diharapkan

menghasilkan *dataset* yang bersih, terstruktur, dan siap untuk dianalisis lebih lanjut, serta mampu memberikan *insight* awal mengenai pola makan sehat, kecukupan gizi, dan risiko gizi buruk. *Dataset* ini diharapkan dapat menjadi alat bantu bagi tenaga kesehatan, pembuat kebijakan, dan peneliti untuk mendukung pengambilan keputusan serta perumusan kebijakan berbasis data.

2. TINJAUAN PUSTAKA

2.1. Gizi Buruk

Gizi buruk adalah keadaan ketika tubuh mengalami kekurangan zat gizi penting dalam jangka waktu tertentu, sehingga berdampak negatif pada fungsi tubuh, imunitas, serta kesehatan secara keseluruhan. Kondisi ini bisa terjadi pada siapa saja, baik anak-anak, remaja, orang dewasa, hingga lansia. Kurangnya asupan energi, protein, vitamin, dan mineral dapat menyebabkan penurunan berat badan drastis, melemahnya daya tahan tubuh, serta gangguan pada pertumbuhan, metabolisme, dan produktivitas seseorang. Jika tidak ditangani, gizi buruk dapat meningkatkan risiko penyakit kronis dan menurunkan kualitas hidup individu [6].

2.2. Pola Makan Sehat

Pola makan sehat merupakan cara mengatur asupan makanan sehari-hari secara seimbang dan teratur agar tubuh memperoleh nutrisi yang cukup untuk menjalankan fungsi vitalnya. Makanan yang dikonsumsi harus mengandung zat gizi lengkap seperti karbohidrat, protein, lemak, vitamin, dan mineral. Tujuannya bukan hanya untuk memenuhi kebutuhan energi, tetapi juga untuk menjaga daya tahan tubuh, mendukung pertumbuhan, serta mencegah berbagai gangguan kesehatan. Pola makan sehat biasanya melibatkan konsumsi beragam jenis makanan, termasuk sumber karbohidrat, lauk berprotein, sayur-mayur, buah-buahan, dan produk susu, serta membatasi makanan yang tinggi gula, garam, dan lemak jenuh [7].

2.3. Data Pipeline

Data *pipeline* adalah serangkaian proses yang saling terhubung, di mana hasil dari satu proses digunakan sebagai *input* untuk proses berikutnya. Dalam konteks pengelolaan data, data *pipeline* berfungsi untuk mengalirkan data dari berbagai sumber yang berbeda ke dalam satu sistem penyimpanan terpusat, seperti data *warehouse*. Tujuannya adalah untuk mempermudah proses pemuatan, pengolahan, dan analisis data secara efisien. Perancangan *pipeline* yang terstruktur dan efisien sangat penting agar kualitas data terjaga dan performa sistem tetap optimal [8].

2.4. Data Ingestion

Data *ingestion* adalah proses awal dalam pengelolaan data yang bertujuan untuk mengumpulkan, mengintegrasikan, dan memindahkan data dari berbagai sumber ke dalam sistem penyimpanan atau platform analitik, seperti data *warehouse* atau data *lake*. Sumber data tersebut bisa berasal dari sistem yang sudah dikenal, seperti *database* rumah sakit, maupun dari perangkat yang belum dikenal seperti IoT atau *wearable device*.

Dalam proses ini, data dapat dikumpulkan secara *batch* (berkala) maupun *streaming (real-time)*, tergantung pada jenis dan sifat sumber data. Selain itu, proses *ingestion* juga sering mencakup pemeriksaan kualitas dan kepercayaan sumber data agar hanya data yang valid dan terpercaya yang digunakan dalam tahap analisis berikutnya [9].

2.5. ETL (Extract, Transform, Load)

ETL (*Extract, Transform, Load*) adalah serangkaian proses penting dalam sistem data *warehouse* yang bertujuan untuk mengambil data dari berbagai sumber, mengolahnya sesuai kebutuhan analisis, dan menyimpannya ke dalam satu tempat penyimpanan terpusat. Proses ini dimulai dengan *extract*, yaitu pengambilan data dari sistem eksternal; dilanjutkan dengan *transform*, yaitu pengolahan atau penyesuaian data agar sesuai dengan struktur dan standar yang dibutuhkan; dan diakhiri dengan *load*, yaitu pemuatan data yang telah diproses ke dalam data *warehouse* [10].

2.6. Transformasi Data

Transformasi data adalah proses mengubah atau menyesuaikan nilai-nilai data agar sesuai dengan kebutuhan analisis dan penyimpanan dalam data *warehouse*. Proses ini bisa meliputi berbagai operasi matematika seperti menambahkan angka konstan, mengalikan nilai, mengkuadratkan atau memangkatkan, mengubah skala data ke bentuk logaritmik, membalik atau memantulkan nilai, mengambil akar kuadrat, bahkan menerapkan fungsi trigonometri seperti gelombang sinus. Tujuannya adalah agar data menjadi lebih mudah dianalisis dan relevan untuk penggunaan tertentu [8].

2.7. Data Storage

Data *storage* atau penyimpanan data adalah media atau perangkat yang digunakan untuk merekam dan menyimpan informasi digital secara permanen atau sementara. Berbeda dengan memori komputer seperti RAM yang hanya menyimpan data sementara dan hilang saat komputer dimatikan, data *storage* memungkinkan data disimpan dengan aman bahkan ketika perangkat dimatikan, sehingga data dapat diakses kembali kapan saja. Penyimpanan data digunakan untuk menyimpan *input* dari pengguna dan hasil *output* komputer, serta mendukung berbagai kebutuhan komputasi modern seperti analisis *big data*, kecerdasan buatan, *machine learning*, dan *Internet of Things* [11].

2.8. Penelitian Terdahulu

“Implementasi *Extract, Transform, Load Process* pada Perancangan Data *Warehouse* Terkait Kualitas Pendidikan di Kabupaten Serang” oleh Raihan Abdul Ghani dan Rudi Kurniawan (2024). Penelitian ini berfokus pada perancangan dan implementasi data *warehouse* dengan menerapkan proses *Extract, Transform, Load* (ETL) untuk mengintegrasikan data pendidikan dari berbagai jenjang sekolah di Kabupaten Serang. Sistem dibangun menggunakan DataPrep by

Trifacta sebagai alat ETL dan Google *BigQuery* sebagai sistem penyimpanan data utama. Tujuan utamanya adalah menciptakan sistem penyimpanan terpusat yang dapat mengatasi permasalahan penyebaran data sekolah di Dapodikdasmen serta meningkatkan efektivitas analisis data pendidikan. Hasil penelitian menunjukkan bahwa penerapan ETL *process* mampu menghasilkan struktur data *warehouse* yang terintegrasi dan efisien, mendukung pengambilan keputusan berbasis data dengan rentang waktu yang relevan. Penelitian ini berkontribusi dalam pengembangan arsitektur data *pipeline* yang adaptif untuk bidang pendidikan, serta menunjukkan efektivitas penerapan ETL pada platform *cloud* sebagai solusi analisis data berskala besar [8].

3. METODE PENELITIAN

3.1. Desain Penelitian

Menggunakan pendekatan rekayasa sistem dalam merancang sebuah data *pipeline* untuk mengolah data pola makan dan menghasilkan analisis yang dapat digunakan sebagai upaya pencegahan gizi buruk di Indonesia. Kerangka penelitian terdiri dari desain *pipeline*, implementasi, evaluasi, dan interpretasi hasil analisis pola makan.

3.2. Sumber Data

Data penelitian ini diperoleh dari sumber data publik, yaitu *Kaggle*, yang berisi informasi mengenai berbagai jenis makanan beserta kandungan gizinya, seperti energi (kalori), protein, karbohidrat, dan lemak. *Dataset* ini dipilih karena bersifat terbuka dan dapat diakses oleh publik, serta mencakup variabel-variabel utama yang penting untuk analisis pola makan. Selain itu, data ini sangat relevan untuk digunakan dalam perancangan sistem rekomendasi nutrisi serta upaya pencegahan gizi buruk.

3.3. Tahapan dalam data *pipeline*

3.3.1. Data *Ingestion*

Tahap pertama adalah proses *data ingestion*, yaitu membaca dan memuat *dataset* ke dalam Google Colab. Proses ini dilakukan menggunakan pustaka *Python* seperti *pandas* dengan fungsi `read_csv()` sehingga data dari sumber publik dapat diakses dan diolah langsung di dalam *notebook*.

3.3.2. Data *Cleaning* dan *Preprocessing*

Setelah data berhasil dimuat, langkah berikutnya adalah *data cleaning & preprocessing*. Pada tahap ini dilakukan pembersihan data dari duplikasi, penanganan nilai yang hilang, serta standarisasi satuan kandungan gizi agar konsisten, misalnya dengan mengonversi energi menjadi per 100 gram makanan. Selain itu, data juga diperiksa terhadap adanya *outlier* yang kemudian ditangani menggunakan metode statistik seperti IQR atau normalisasi sesuai kebutuhan, sehingga kualitas data dapat tetap terjaga.

3.3.3. Data *Transformation* dan *Feature Engineering*

Pada bagian ini, data diolah lebih lanjut untuk menghasilkan informasi yang lebih bermakna, misalnya dengan menambahkan variabel turunan seperti total energi dan rasio antara protein, karbohidrat, serta lemak. Selain itu, dilakukan pula proses kategorisasi makanan berdasarkan tingkat kesehatannya sesuai dengan standar Angka Kecukupan Gizi (AKG) atau rekomendasi WHO. Apabila diperlukan, normalisasi variabel juga dilakukan agar hasil analisis lebih akurat dan dapat dibandingkan dengan baik.

3.3.4. Data *Storage*

Setelah data melalui proses pembersihan dan transformasi, hasil akhirnya disimpan kembali dalam format terstruktur, seperti CSV. Penyimpanan ini bertujuan agar *dataset* yang sudah siap analisis dapat digunakan ulang pada tahap selanjutnya maupun untuk penelitian lain di kemudian hari.

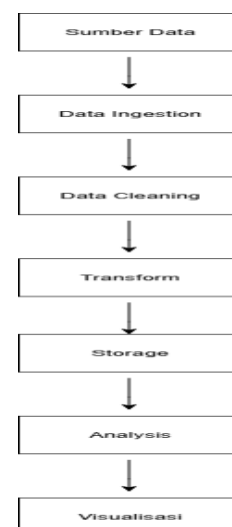
3.3.5. Data *Analysis* dan *Visualization*

Analisis deskriptif digunakan untuk menggambarkan distribusi variabel utama seperti kalori, protein, karbohidrat, dan lemak. Selain itu, dilakukan analisis hubungan antar variabel gizi, misalnya untuk melihat korelasi antara asupan energi dengan status gizi. Hasil analisis ini divisualisasikan menggunakan pustaka *matplotlib* dan *seaborn* dalam bentuk grafik seperti histogram, *boxplot*, *bar chart*, dan *heatmap* korelasi, sehingga pola konsumsi makanan dapat lebih mudah dipahami.

3.4. Alat dan Teknologi

Penelitian ini menggunakan *Python* di Google Colab sebagai lingkungan pemrograman karena mudah diakses dan mendukung komputasi berbasis *cloud*. Pustaka yang dipakai antara lain *pandas* dan *numpy* untuk pengolahan data, *matplotlib* dan *seaborn* untuk visualisasi, serta *scipy* dan *scikit-learn* untuk analisis statistik maupun *clustering*.

3.5. Arsitektur Sistem



Gambar 1. Flowchart Arsitektur Data Pipeline untuk Analisis Pola Makan Sehat

Gambar 1 menunjukkan arsitektur data *pipeline* yang terdiri dari beberapa tahap utama, yaitu: sumber data, data *ingestion*, data *cleaning*, transformasi, penyimpanan (*storage*), analisis, dan visualisasi. *Pipeline* ini dirancang untuk mengolah data pola makan dari berbagai sumber agar dapat disimpan dalam format terstruktur (CSV) dan siap digunakan untuk analisis lanjutan. Dalam implementasi saat ini, *pipeline* dibangun hingga tahap penyimpanan, namun tetap terbuka untuk pengembangan ke tahap analisis dan visualisasi di masa mendatang.

4. HASIL PEMBAHASAN

4.1. Hasil Implementasi Data Pipeline

4.1.1. Data Ingestion

```
# === STEP 1: IMPORT & EXTRACT ===
import pandas as pd
import numpy as np

# Load dataset
meals = pd.read_csv("detailed_meals_macros.csv")
nutrition = pd.read_csv("Food_and_Nutrition_.csv")
```

Gambar 2. Proses Load Data

Dataset yang digunakan diimpor dari *Kaggle*, dengan dua sumber utama yaitu *detailed_meals_macros.csv* dan *Food_and_Nutrition.csv*. *Dataset* ini berisi informasi detail mengenai komposisi makanan serta data individu yang terkait dengan pola makan. Proses *import* berhasil dilakukan, dan kolom pada kedua *dataset* telah dinormalisasi agar konsisten untuk tahap pengolahan selanjutnya.

4.1.2. Data Transformation

```
# === STEP 2: TRANSFORM (Gabung Data) ===
# Tentukan key gabungan
merge_keys = ["Ages", "Gender", "Height", "Weight", "Activity_Level", "Dietary_Preference"]

# Pilih subset kolom penting dari nutrition
menu_cols = ["Breakfast_Suggestion", "Lunch_Suggestion", "Dinner_Suggestion", "Snack_Suggestion"]
nutrition_subset = nutrition[merge_keys + menu_cols]

# Gabungkan dataset
combined = pd.merge(meals, nutrition_subset, on=merge_keys, how="left")
print("Data berhasil digabung, shape:", combined.shape)
```

Gambar 3. Proses Transform (Penggabungan Data)

Dataset Meals dan *Nutrition* digabung menggunakan key tertentu (*Ages*, *Gender*, *Height*, *Weight*, *Activity_Level*, *Dietary_Preference*). Hasil penggabungan menghasilkan *dataset* baru dengan 2.514 baris dan 39 kolom, yang memuat data konsumsi harian individu beserta kandungan nutrisi dari menu yang disarankan.

4.1.3. Perhitungan Total Nutrisi per Individu

```
# === STEP 3: HITUNG TOTAL NUTRISI PER INDIVIDU/HARI ===
# Kolom nutrisi utama yang tersedia di dataset
nutrient_cols = ["Calories", "Protein", "Carbohydrates", "Fiber", "Fat"]

# Kelompokkan per individu berdasarkan merge_keys
daily_totals = combined.groupby(merge_keys)[nutrient_cols].sum().reset_index()
print("Total nutrisi per individu sudah dihitung")
print(daily_totals.head())
```

Ages	Gender	Height	Weight	Activity_Level	Dietary_Preference	Calories	Protein	Carbohydrates	Fiber	Fat	
0	18	Female	152	56	Moderately Active	Vegetarian	1949	146	195	23.48	65
1	18	Female	160	53	Moderately Active	Vegan	1784	119	219	26.28	48
2	18	Female	165	54	Extremely Active	Vegetarian	1994	150	200	24.00	66
3	18	Female	168	75	Extremely Active	Vegetarian	2930	220	204	35.28	98
4	18	Female	170	52	Very Active	Pescatarian	1983	115	257	30.84	55

Gambar 4. Perhitungan Total Nutrisi per Individu

Pada tahap ini dilakukan perhitungan total asupan nutrisi per individu per hari, mencakup energi (kalori), protein, karbohidrat, serat, dan lemak. Perhitungan dilakukan dengan menjumlahkan seluruh menu makanan yang dikonsumsi oleh masing-masing individu berdasarkan atribut kunci (usia, jenis kelamin, tinggi, berat, tingkat aktivitas, dan preferensi diet).

Hasil yang diperoleh menunjukkan bahwa setiap individu memiliki total asupan nutrisi yang berbeda-beda. Sebagai contoh, individu dengan tingkat aktivitas sangat tinggi memiliki asupan kalori harian yang lebih besar dibandingkan individu dengan tingkat aktivitas sedang. Begitu juga, preferensi diet seperti vegetarian atau *vegan* berpengaruh terhadap variasi asupan protein dan lemak.

4.1.4. Perbandingan Dengan Standar Gizi

```
# === STEP 4: BANDINGKAN DENGAN STANDAR ===
# Standar harian sederhana (bisa diganti tabel resmi Kemenkes/WHO)
standards = {
    "Calories": (1800, 2200),
    "Protein": (50, 70),
    "Carbohydrates": (225, 325),
    "Fiber": (25, 35),
    "Fat": (50, 80)
}

# Fungsi untuk menilai status nutrisi
def assess_nutrient(value, rec_min, rec_max):
    if pd.isna(value):
        return "Missing"
    if value < rec_min:
        return "Kurang"
    elif value > rec_max:
        return "Lebih"
    else:
        return "Cukup"
```

Gambar 5. Proses Membandingkan dengan Standar Gizi

Pada tahap ini, total asupan nutrisi individu dibandingkan dengan standar kecukupan gizi harian yang direkomendasikan oleh WHO/Kemenkes, yaitu kalori (1800–2200 kkal), protein (50–70 g), karbohidrat (225–325 g), serat (25–35 g), dan lemak (50–80 g). Implementasi dilakukan dengan membandingkan nilai aktual terhadap ambang batas tersebut.

4.1.5. Ringkasan Status Gizi per Individu

```
# Hitung proporsi
summary["Total_Nutrients"] = summary[["Cukup", "Kurang", "Lebih", "Missing"]].sum(axis=1)
summary["Pct_Cukup"] = summary["Cukup"] / summary["Total_Nutrients"]
summary["Pct_Kurang"] = summary["Kurang"] / summary["Total_Nutrients"]

# Label keseluruhan
def overall_label(row):
    if row["Pct_Kurang"] > 0.5:
        return "Kurang Sehat"
    elif row["Pct_Cukup"] >= 0.6:
        return "Seimbang"
    else:
        return "Perlu Perbaikan"

summary["Overall_Label"] = summary.apply(overall_label, axis=1)
```

Gambar 6. Membuat Label Status Gizi

Distribusi Status Pola Makan Harian:

Overall_Label	Count
Perlu Perbaikan	936
Seimbang	319
Kurang Sehat	315

Persentase (%)	Overall_Label	Persentase (%)
Overall_Label	Perlu Perbaikan	59.617834
	Seimbang	20.318471
	Kurang Sehat	20.063694

Gambar 7. Hasil Distribusi dan Persentase Pola Makan Harian

Setelah dilakukan perbandingan per nutrisi, sistem membuat ringkasan status gizi per individu. Proses ini mengelompokkan hasil menjadi *cukup*, *kurang*, atau *lebih*. Selanjutnya, dilakukan klasifikasi *overall* dengan ketentuan:

- “Kurang Sehat” jika lebih dari 50% nutrisi tidak tercukupi,
- “Seimbang” jika lebih dari 60% nutrisi tercukupi,
- “Perlu Perbaikan” untuk kasus lainnya.

Hasil analisis menunjukkan mayoritas responden berada pada kategori “Perlu Perbaikan” (59,6%), disusul oleh “Seimbang” (20,3%), dan “Kurang Sehat” (19,9%). Hal ini mengindikasikan bahwa pola makan sebagian besar responden belum memenuhi standar gizi seimbang.

4.1.6. Ekspor dan Penyimpanan Hasil Analisis

```
# Simpan hasil
summary.to_csv("daily_nutrition_summary.csv", index=False)
long_df.to_csv("daily_nutrition_long.csv", index=False)

print("\n✅ Hasil analisis disimpan ke CSV")
```

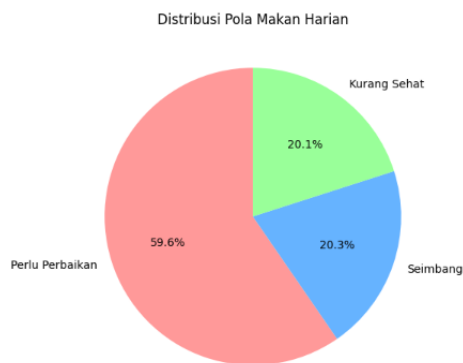
✅ Hasil analisis disimpan ke CSV

Gambar 7. Penyimpanan Hasil kedalam Bentuk CSV

Hasil analisis disimpan dalam format **CSV** untuk memudahkan dokumentasi dan pemanfaatan data lebih lanjut. File `daily_nutrition_summary.csv` berisi ringkasan status nutrisi per individu, sedangkan `daily_nutrition_long.csv` menyimpan detail penilaian setiap jenis nutrisi terhadap standar gizi.

4.2. Visualisasi dan Analisis Pola Makan

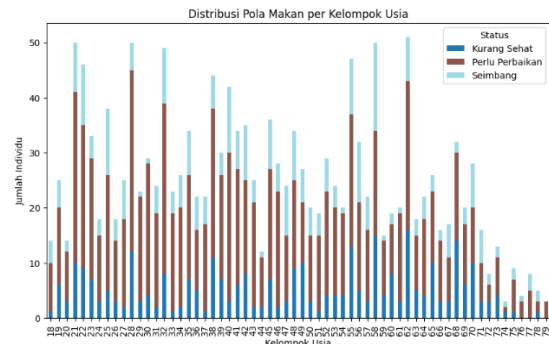
4.2.1 Distribusi Pola Makan Harian



Gambar 8. Visualisasi Pola Makan Harian

Grafik *pie chart* menunjukkan distribusi umum status pola makan seluruh individu. Hasilnya, sebesar 59,6% responden memiliki pola makan yang perlu perbaikan, 20,3% seimbang, dan 20,1% kurang sehat. Temuan ini mengindikasikan bahwa sebagian besar individu masih belum mencapai keseimbangan gizi harian, sehingga perlu adanya edukasi dan perbaikan pola konsumsi.

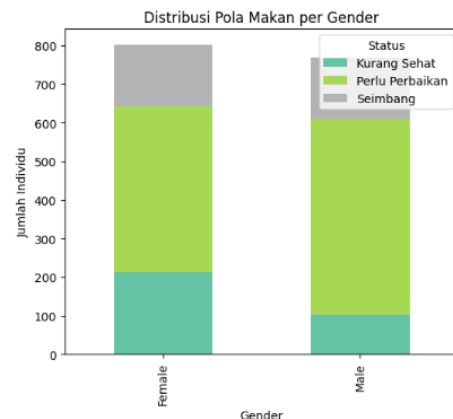
4.2.2 Distribusi Pola Makan Berdasarkan Usia



Gambar 9. Visualisasi Pola Makan Berdasarkan Usia

Grafik batang (*bar chart*) memperlihatkan variasi pola makan berdasarkan kelompok usia. Tampak bahwa kelompok usia produktif (20–50 tahun) mendominasi kategori “Perlu Perbaikan”, sedangkan pola makan seimbang lebih banyak ditemukan pada usia menengah ke atas. Hal ini menunjukkan bahwa gaya hidup dan aktivitas pada usia produktif berpengaruh terhadap ketidakseimbangan pola makan.

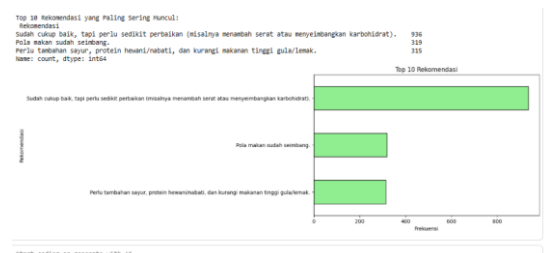
4.2.3 Distribusi Pola Makan Berdasarkan Jenis Kelamin



Gambar 10. Visualisasi Pola Makan Berdasarkan Gender

Hasil menunjukkan bahwa baik laki-laki maupun perempuan sebagian besar masuk kategori “Perlu Perbaikan”, namun proporsinya sedikit lebih tinggi pada perempuan. Perbedaan ini dapat disebabkan oleh faktor preferensi makanan dan tingkat aktivitas harian yang berbeda antara gender.

4.2.4 Rekomendasi Pola makan



Start coding or generate with AI.

Gambar 11. Rekomendasi Pola Makan Berdasarkan Hasil Analisis

Dari grafik terlihat bahwa rekomendasi yang paling banyak diberikan adalah “Sudah cukup baik, tapi perlu sedikit perbaikan (misalnya menambah serat atau menyeimbangkan karbohidrat)” dengan jumlah 936 responden. Hal ini menunjukkan bahwa sebagian besar masyarakat sudah memiliki pola makan yang relatif baik, namun masih perlu sedikit penyesuaian agar lebih seimbang sesuai pedoman gizi seimbang.

Rekomendasi terbanyak kedua adalah “Pola makan sudah seimbang” (319 responden), sedangkan rekomendasi ketiga yang sering muncul adalah “Perlu tambahan sayur, protein hewani/nabati, dan kurangi makanan tinggi gula/lemak” (315 responden). Temuan ini mengindikasikan bahwa sebagian masyarakat masih kekurangan konsumsi sayur dan protein serta cenderung mengonsumsi makanan tinggi gula dan lemak.

Secara keseluruhan, hasil ini menggambarkan bahwa tingkat kesadaran gizi masyarakat cukup baik, namun masih diperlukan edukasi lanjutan mengenai keseimbangan gizi dan pentingnya memperbanyak konsumsi serat, protein, serta mengurangi makanan olahan tinggi gula dan lemak.

4.3. Diskusi Temuan Utama

Hasil analisis menunjukkan bahwa sebagian besar individu ($\pm 59,6\%$) memiliki pola makan yang perlu perbaikan, dengan 20,3% sudah seimbang dan 20,1% tergolong kurang sehat. Temuan ini menunjukkan bahwa tingkat kesadaran gizi masyarakat masih perlu ditingkatkan.

Distribusi berdasarkan usia memperlihatkan bahwa kelompok usia produktif (20–50 tahun) cenderung memiliki pola makan yang tidak seimbang. Hal ini dapat disebabkan oleh gaya hidup cepat saji, waktu makan yang tidak teratur, serta kebiasaan konsumsi tinggi karbohidrat dan lemak.

Sementara itu, analisis berdasarkan gender menunjukkan bahwa baik laki-laki maupun perempuan sama-sama mendominasi kategori “Perlu Perbaikan”. Namun, sedikit lebih banyak perempuan yang memiliki pola makan seimbang, kemungkinan karena lebih memperhatikan asupan makanan dan kalori harian.

Dari sisi rekomendasi sistem, hasil menunjukkan bahwa sebagian besar individu sudah cukup baik namun masih perlu penyesuaian pada asupan serat, protein, dan karbohidrat agar memenuhi standar gizi seimbang Kemenkes/WHO.

Secara keseluruhan, *pipeline* yang dibangun berhasil memberikan gambaran komprehensif tentang pola makan masyarakat dan dapat menjadi dasar untuk pengambilan keputusan kebijakan gizi atau pengembangan aplikasi rekomendasi makanan sehat.

4.4. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan.

1. Data yang digunakan berasal dari *dataset* publik di *Kaggle*, sehingga validitas dan representatif data sangat bergantung pada kualitas dan kelengkapan data yang telah disediakan oleh sumber tersebut. Data mungkin tidak sepenuhnya mencerminkan

kondisi populasi sebenarnya karena adanya potensi bias pada proses pengumpulan data.

2. Analisis dilakukan dengan pendekatan deskriptif berbasis *Python* tanpa melakukan uji statistik lanjutan atau pemodelan prediktif. Hal ini membatasi kemampuan penelitian untuk menarik kesimpulan kausal antara variabel-variabel seperti asupan gizi, kebiasaan makan, dan rekomendasi pola makan.
3. Variabel-variabel yang tersedia dalam *dataset* masih terbatas dan tidak mencakup faktor-faktor lain yang dapat memengaruhi pola makan, seperti aktivitas fisik, tingkat ekonomi, atau kondisi kesehatan individu.
4. Rekomendasi yang dihasilkan bersifat umum berdasarkan tren data dan tidak dipersonalisasi sesuai kebutuhan individu. Oleh karena itu, hasil penelitian ini sebaiknya digunakan sebagai gambaran umum pola konsumsi gizi, bukan sebagai acuan medis atau diet pribadi.

5. KESIMPULAN

Penelitian ini berhasil merancang dan mengimplementasikan data *pipeline* untuk menganalisis pola makan sehat masyarakat Indonesia berdasarkan data publik dari *Kaggle*. *Pipeline* yang dibangun terdiri dari enam tahapan utama, yaitu pengumpulan data, pembersihan, transformasi, analisis, visualisasi, dan penyimpanan hasil. Seluruh tahapan telah berjalan dengan baik menggunakan bahasa pemrograman *Python*, menghasilkan data yang terstruktur dan siap untuk analisis lanjutan.

Hasil analisis menunjukkan bahwa sebagian besar individu ($\pm 59,6\%$) memiliki pola makan yang perlu perbaikan, sementara hanya sebagian kecil yang sudah menerapkan pola makan seimbang ($\pm 20,3\%$). Faktor usia dan jenis kelamin juga menunjukkan variasi, di mana kelompok usia produktif cenderung memiliki pola makan yang kurang seimbang. Rekomendasi sistem menunjukkan bahwa mayoritas masyarakat hanya memerlukan sedikit penyesuaian, terutama pada konsumsi serat, protein, dan pengurangan makanan tinggi gula dan lemak.

Pipeline ini memberikan gambaran menyeluruh mengenai kondisi pola makan masyarakat dan dapat dijadikan alat bantu analisis bagi lembaga kesehatan, peneliti, maupun pembuat kebijakan dalam merumuskan strategi pencegahan gizi buruk. Dengan penerapan yang lebih luas dan data yang lebih beragam, *pipeline* ini berpotensi menjadi fondasi sistem pemantauan gizi nasional berbasis data.

DAFTAR PUSTAKA

- [1] A. M. "SSGI 2024: Prevalensi Stunting Nasional Turun Menjadi 19,8%," 26 Mei 2025. [Online]. Available: <https://kemkes.go.id/id/ssgi-2024-prevalensi->

- stunting-nasional-turun-menjadi-198.
[Accessed 30 September 2025].
- [2] "SURVEI KESEHATAN INDONESIA (SKI) 2023," in *SURVEI KESEHATAN INDONESIA DALAM ANGKA*, Kemenkes BKPK, 2023.
- [3] BPN, "Gugah Kesadaran Masyarakat Pentingnya Sayur dan Buah Demi Capai Pola Makan Bergizi Seimbang," 20 Oktober 2024. [Online]. Available: <https://badanpangan.go.id/blog/post/gugah-kesadaran-masyarakat-pentingnya-sayur-dan-buah-demi-capai-pola-makan-bergizi-seimbang>. [Accessed 30 September 2025].
- [4] N. W. Panjaitan, S. S. Hasibuan and F. , "Analisis Tingkat Pengetahuan Tentang Gizi Seimbang Bagi Remaja di MTS Ex PGA Proyek UNIVA Medan," *PubHealth Jurnal Kesehatan Masyarakat*, vol. 1, no. 1, 2022.
- [5] M. T. Hidayat, C. Y. A. M. and A. I. Maula, "IMPLEMENTASI DASHBOARD BUSINESS INTELLIGENCE DENGAN QLIK SENSE UNTUK ANALISIS NUTRISI MAKANAN," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 13, no. 1, 2025.
- [6] R. A. Ghani and R. Kurniawan, "IMPLEMENTASI EXTRACT, TRANSFORM, LOAD PROCESS PADA PERANCANGAN DATA WAREHOUSE TERKAIT KUALITAS PENDIDIKAN DI KABUPATEN SERANG," *Jurnal Mahasiswa Teknik Informatika*, vol. 8, no. 2, pp. 2083-2090, April 2024.
- [7] A. Mavrogiorgou, t. Kiourtis, G. Manias, h. Symvoulidis and D. Kyriazis, "Batch and Streaming Data Ingestion towards Creating Holistic Health Records," *EmergingScience Journal*, vol. 7, no. 2, April 2023.
- [8] L. D. A. Oktafiani, D. Rokhmah, K. and Z. E. P. Arini, "Edukasi Pencegahan Gizi Buruk Pada Ibu Balita dan Kader Posyandu di Desa Kajar Kabupaten Bondowoso," *Journal of Community Development*, vol. 4, no. 3, pp. 263-270, April 2024.
- [9] W. A. Wicaksana, G. Wulandari and D. S. Y. Kartika, "Mewujudkan Perilaku Hidup Bersih Dan Sehat Dengan Cuci Tangan, Pola Makan Sehat, Dan Pengelolaan Sampah," *Pandawa : Pusat Publikasi Hasil Pengabdian Masyarakat*, vol. 2, no. 1, pp. 143-151, January 2024.
- [10] P. Sari, L. I. Kesuma, M. Afrina and D. Kurniawan, "Pemodelan Integrasi Data Barang Milik Negara di Perguruan Tinggi Menggunakan Metode ETL (Extract, Transform, Load) dengan Pentaho," *The Indonesian Journal of Computer Science*, vol. 13, no. 5, pp. 8410-8425, 5 October 2024.
- [1] IBM, "Apa yang dimaksud dengan penyimpanan data?," 2024. [Online]. Available: [id/think/topics/data-storage. \[Accessed 07 Oktober 2025\].](https://www.ibm.com/id-</p>
</div>
<div data-bbox=)