



PENERAPAN ALGORITMA K-MEANS CLUSTERING DATA GIZI BALITA PADA UPTD PUSKESMAS

Abi juliansyah¹, Siska Maurani^{2*}, Nurrohim³, Wardianto⁴

^{1,3}Informaika, Sains dan Teknologi, Universitas Nurul Huda, Indonesia

Email: siskamaurani3@gmail.com²

Informasi Artikel

Riwayat Artikel:

Diterima: Tanggal-Bulan-Tahun

Direvisi: Tanggal-Bulan-Tahun

Diterbitkan: Tanggal-Bulan-Tahun

Keywords:

Toddler nutritional status,

Mid-Upper Arm Circumference (MUAC),

Python,

Data analysis

Data cleaning

ABSTRAK

The nutritional status of toddlers is one of the important indicators used to assess children's health and growth. Determining nutritional status manually requires a relatively long processing time and is prone to errors; therefore, the utilization of technology is necessary to improve the efficiency of data processing. This study aims to analyze the nutritional status of toddlers based on the Mid-Upper Arm Circumference (MUAC) indicator using the Python programming language. The research method consisted of three stages: dataset exploration, data cleaning, and nutritional status analysis. The dataset exploration stage was conducted to identify dataset characteristics through data structure inspection, descriptive statistics, missing value detection, duplicate data identification, histograms, and boxplots. Subsequently, the data cleaning process included duplicate removal, missing value handling, data type validation, and anthropometric value validation to produce a dataset suitable for analysis. The analysis stage classified toddlers' nutritional status based on MUAC values into four categories: Severe Malnutrition, Undernutrition, Normal Nutrition, and Overnutrition using a classification function implemented in Python. The results showed that the dataset exploration and data cleaning processes successfully produced a dataset ready for analysis, while the classification process automatically categorized all toddler data according to their nutritional status based on MUAC values. The use of Python proved effective in accelerating data processing, improving the consistency of classification results, and reducing potential errors in determining toddlers' nutritional status.

Kata Kunci:

Kata Kunci 1,

Kata Kunci 2,

Kata Kunci 3,

ABSTRAK

Status gizi balita merupakan salah satu indikator penting untuk menilai kesehatan dan pertumbuhan anak. Penentuan status gizi secara manual memerlukan waktu pemrosesan yang relatif lama dan rentan terhadap kesalahan; oleh karena itu, pemanfaatan teknologi diperlukan untuk meningkatkan efisiensi pemrosesan data. Penelitian ini bertujuan untuk menganalisis status gizi balita berdasarkan indikator Lingkar Lengan Atas (LiLA) menggunakan bahasa pemrograman Python. Metode penelitian terdiri dari tiga tahapan: eksplorasi dataset, pembersihan data, dan analisis status gizi. Tahap eksplorasi dataset dilakukan untuk mengidentifikasi karakteristik dataset melalui pemeriksaan struktur data, statistik deskriptif, deteksi nilai yang hilang (*missing value*), identifikasi data duplikat, histogram, dan *boxplot*. Selanjutnya, proses pembersihan data mencakup penghapusan data duplikat, penanganan nilai yang hilang, validasi tipe data, dan validasi nilai antropometri untuk menghasilkan dataset yang layak dianalisis. Tahap analisis mengklasifikasikan status gizi balita berdasarkan nilai LiLA ke dalam empat kategori: Gizi Buruk, Gizi Kurang, Gizi Baik, dan Gizi Lebih, menggunakan fungsi klasifikasi yang diimplementasikan dalam Python. Hasil penelitian menunjukkan bahwa proses eksplorasi dan pembersihan data berhasil menghasilkan dataset yang siap dianalisis, sementara proses klasifikasi secara otomatis mengategorikan seluruh data balita sesuai dengan status gizinya berdasarkan nilai LiLA. Penggunaan Python terbukti efektif dalam mempercepat pemrosesan data, meningkatkan konsistensi hasil klasifikasi, serta mengurangi potensi kesalahan dalam penentuan status gizi balita.

Penulis Korespondensi:

Nama Penulis,
Program Studi....,
Universitas.....
Email: xxxx@xxxxx.xxx

Artikel ini merupakan artikel akses terbuka di bawah lisensi [CC BY-SA](#).

**1. PENDAHULUAN**

Kemajuan teknologi di sektor kesehatan saat ini berlangsung sangat pesat, mendorong pengembangan aplikasi yang membantu menghemat waktu, biaya, dan tenaga dalam berbagai aspek aktivitas manusia. Teknologi ini juga memfasilitasi perolehan data berkualitas tinggi—seperti data yang diperlukan untuk mengelola informasi gizi balita—sehingga memperlancar proses pengolahan data.

Di sektor kesehatan, kemajuan teknologi informasi telah secara signifikan meningkatkan kualitas layanan. Hal ini sangat krusial dalam bidang medis, di mana pengetahuan berkembang dengan cepat—dengan sekitar 750.000 artikel baru diterbitkan di jurnal medis setiap tahunnya—sehingga dokter berisiko tertinggal jika tidak memanfaatkan sarana untuk terus mengikuti perkembangan terkini. Selain itu, teknologi informasi memiliki kemampuan untuk menyaring data dan mengubahnya menjadi informasi yang dapat ditindaklanjuti, seperti kategorisasi status gizi balita [1].

K-means clustering adalah metode pengelompokan data yang membagi data ke dalam satu atau lebih kluster. Titik-titik data yang memiliki karakteristik serupa dikelompokkan bersama, sedangkan data dengan karakteristik berbeda ditempatkan dalam kluster terpisah, sehingga meminimalkan variasi di dalam setiap kluster [2].

Ni Komang Sri Julyantari melakukan penelitian berjudul "K-Means untuk Pengelompokan Status Gizi Balita (Studi Kasus: Banjar Titih)." Penelitian ini menggunakan data status gizi balita dengan melibatkan atribut seperti usia, jenis kelamin, dan berat badan; atribut-atribut ini menjadi input yang diproses untuk menghasilkan output berupa kelompok atau kluster yang berbeda. Sampel data diperoleh dari tiga Pos Pelayanan Terpadu (Posyandu)—Titih Kelod, Titih Tengah, dan Titih Kaler—pada tanggal 21 Agustus 2021, yang mencakup total 46 balita. Penelitian ini menerapkan proses klasifikasi status gizi balita dan mengidentifikasi tiga kluster: Gizi Kurang, Normal, dan Gizi Lebih. Hasil pengelompokan status gizi di Banjar Titih menunjukkan adanya tiga kluster: 30,43% dengan gizi kurang, 47,83% dengan gizi normal, dan 21,74% dengan gizi lebih; yang perlu dicatat, terdapat angka yang signifikan baik untuk kategori gizi kurang maupun gizi lebih [3].

Abdi Subayu melakukan penelitian berjudul "Analisis K-Means terhadap Stunting Gizi pada Balita: Sebuah Tinjauan Sistematis." Penelitian ini menggunakan teknik *data mining*—khususnya algoritma K-Means—untuk pengelompokan (*clustering*) data. Algoritma K-Means dipilih karena kesederhanaannya. Pengumpulan data berfokus pada tiga parameter balita: tinggi badan, berat badan, dan lingkaran kepala. Studi mengenai status gizi ini dilakukan di Kota Tasikmalaya, sebuah wilayah dengan jumlah balita terdampak *stunting* mencapai 5.290 jiwa. Hasil pengelompokan menggunakan algoritma K-Means menunjukkan bahwa 17 titik data ditempatkan pada kelompok yang tepat, sehingga menghasilkan tingkat akurasi hanya sebesar 34% ($17/50 = 43\%$). Dengan demikian, dapat disimpulkan bahwa algoritma K-Means tidak cocok digunakan untuk mengelompokkan status gizi balita [4].

Kurang gizi merupakan masalah kesehatan yang sering dialami oleh balita di Indonesia. Berdasarkan survei Pemantauan Status Gizi (PSG) tahun 2017 yang dilakukan oleh Kementerian Kesehatan, prevalensi masalah gizi pada anak di bawah usia lima tahun (balita) mencapai 17,8% pada tahun 2017. Angka tersebut mencakup balita yang mengalami gizi buruk (3,8%) dan gizi kurang (14%), sebagaimana ditunjukkan oleh data mengenai masalah gizi balita di Indonesia pada periode 2016 hingga 2017 [5].

Kurang gizi atau malnutrisi pada anak balita berdampak negatif terhadap pertumbuhan fisik maupun mental, serta menghambat pencapaian tahapan perkembangan seperti kemampuan berbicara, berjalan, dan makan. Keterlibatan orang tua dan petugas Pusat Kesehatan Masyarakat (Puskesmas) sangat penting untuk memantau status gizi balita, karena malnutrisi tidak muncul secara tiba-tiba seperti penyakit pada umumnya; tanda-tanda seperti berat badan di bawah standar dan *stunting* dapat menjadi indikator awal [6]. Kegagalan dalam menyediakan makanan bergizi bagi balita dapat menyebabkan kurang gizi; status gizi balita sangat erat kaitannya dengan orang tua, karena hal ini menentukan kualitas keturunan keluarga tersebut [7].

Masa balita merupakan periode kritis, karena proses pertumbuhan dan perkembangan pada masa ini meletakkan dasar bagi keberhasilan tahapan perkembangan selanjutnya [8]. Status gizi dipengaruhi oleh asupan makanan dan pemanfaatan zat gizi oleh tubuh. Ketika tubuh menerima zat gizi yang memadai dan memanfaatkannya secara efisien, status gizi optimal akan tercapai, sehingga memungkinkan pertumbuhan fisik, perkembangan otak, kapasitas kerja, dan kesehatan umum yang maksimal [9].

Pengelompokan populasi berdasarkan kriteria status gizi—seperti gizi buruk, kurang gizi, gizi baik, kelebihan berat badan, dan obesitas—memudahkan perancangan dan pelaksanaan program gizi yang tepat sasaran oleh para pembuat kebijakan dan tenaga kesehatan [10].

Malnutrisi yang terjadi selama periode kritis pertumbuhan dan perkembangan ini dapat menimbulkan konsekuensi jangka panjang; perubahan yang terjadi selama "1.000 Hari Pertama Kehidupan" (HPK) bersifat permanen. Gizi yang baik pada periode ini mendukung terciptanya sumber daya manusia yang berkualitas dengan memastikan pertumbuhan linier yang optimal (mencegah *stunting*), perkembangan kognitif (kecerdasan) yang optimal, serta risiko penyakit tidak menular yang lebih rendah [11].

Mengingat pentingnya pertumbuhan dan perkembangan selama 1.000 hari pertama kehidupan, pemantauan pertumbuhan anak balita harus dilakukan secara optimal. Pemantauan tersebut merupakan bagian dari standar pelayanan minimal di tingkat daerah dan berfungsi menyediakan data yang diperlukan untuk menilai sejauh mana masalah gizi terjadi di suatu wilayah [12].

Selain pola makan, pengetahuan ibu mengenai gizi merupakan faktor penting lainnya. Sebagai pengasuh utama, ibu memegang peranan kunci dalam memilih, menyiapkan, dan menyajikan makanan bagi anak. Pengetahuan yang memadai mengenai zat gizi, frekuensi makan, dan pemberian makanan pendamping ASI memengaruhi kualitas asupan makanan anak [13].

Perkembangan anak mencerminkan perubahan kualitatif dalam kemampuan fungsional yang muncul dari interaksi antara pematangan sistem saraf pusat dan organ-organ yang memengaruhinya—suatu proses yang melibatkan pertumbuhan, pengendalian, dan pembelajaran yang berkelanjutan (Istiqomah dkk., 2024). Aspek perkembangan anak mencakup keterampilan motorik kasar dan halus, penguasaan bahasa, serta kemandirian dan kemampuan sosial; hal-hal ini dipengaruhi oleh faktor-faktor seperti upaya stimulasi, praktik pengasuhan, latar belakang sosial-ekonomi, pendidikan dan pengetahuan ibu, pemberian ASI eksklusif, lingkungan sekitar, kekurangan zat besi, dan gizi (Avisha dkk., 2023). Gizi merupakan salah satu faktor penentu utama perkembangan. Asupan energi dan zat gizi yang memadai berperan penting dalam mendukung perkembangan fisik, kognitif, dan sosial-emosional anak usia balita [14].

2. METODE PENELITIAN

Penelitian ini menggunakan metode analisis data kuantitatif dengan pendekatan deskriptif. Metode ini bertujuan untuk menganalisis status gizi balita berdasarkan data antropometri yang diperoleh dari dataset penelitian. Analisis dilakukan secara bertahap, mulai dari eksplorasi, pembersihan, dan validasi data hingga klasifikasi status gizi berdasarkan nilai Lingkar Lengan Atas (LiLA/MUAC). Seluruh pemrosesan data dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab, guna memastikan proses analisis yang sistematis, cepat, dan dapat direproduksi [15].

Penelitian dimulai dengan memuat dataset balita yang berformat Microsoft Excel (.xlsx). Selanjutnya, dilakukan eksplorasi data untuk memahami karakteristik dataset, termasuk struktur, tipe data atribut, statistik deskriptif, pemeriksaan nilai yang hilang (*missing values*) dan duplikat, serta visualisasi distribusi data menggunakan histogram dan *boxplot*. Tahap berikutnya melibatkan pembersihan data untuk menghasilkan dataset yang telah disempurnakan melalui penghapusan data duplikat dan rekaman dengan nilai kosong (*null*), validasi tipe data numerik serta rentang nilai antropometri, dan pengindeksan ulang data.

Setelah dataset yang bersih diperoleh, penelitian dilanjutkan dengan mengklasifikasikan status gizi balita berdasarkan nilai LiLA mereka. Setiap balita dikategorikan ke dalam salah satu dari empat kelompok status gizi—Gizi Buruk (*Severe Malnutrition*), Gizi Kurang/Sedang (*Underweight/Moderate Malnutrition*), Gizi Baik, dan Gizi Lebih (*Overnutrition*)—sesuai dengan interval nilai LiLA yang telah ditetapkan sebelumnya. Hasil klasifikasi kemudian divisualisasikan dalam bentuk tabel dan grafik untuk mempermudah analisis serta interpretasi temuan.

2.2 Dataset Penelitian

Dataset yang digunakan pada penelitian ini merupakan data antropometri balita yang disimpan dalam format Microsoft Excel (.xlsx). Setiap baris pada dataset merepresentasikan satu data balita, sedangkan setiap kolom berisi atribut yang digunakan dalam proses analisis status gizi. Dataset ini merupakan data identitas balita dan data pengukuran antropometri yang digunakan sebagai dasar klasifikasi status gizi. Atribut yang digunakan pada penelitian ini disajikan pada Tabel 3.1

No	Nama Atribut	Tipe Data	Keterangan
1	ID	Integer	Identitas balita
2	Nama_Balita	String	Nama balita
3	Jenis_Kelamin	String	Jenis kelamin balita
4	Umur_Bulan	Integer	Umur balita (bulan)
5	Berat_Badan_kg	Float	Berat badan (kg)

6	Tinggi_Badan_cm	Float	Tinggi badan (cm)
7	Lingkar_Lengan_cm	Float	Lingkar Lengan Atas (cm)

Dataset tersebut kemudian diolah menggunakan Python untuk memperoleh informasi mengenai karakteristik data, membersihkan data yang tidak valid, serta melakukan analisis status gizi balita.

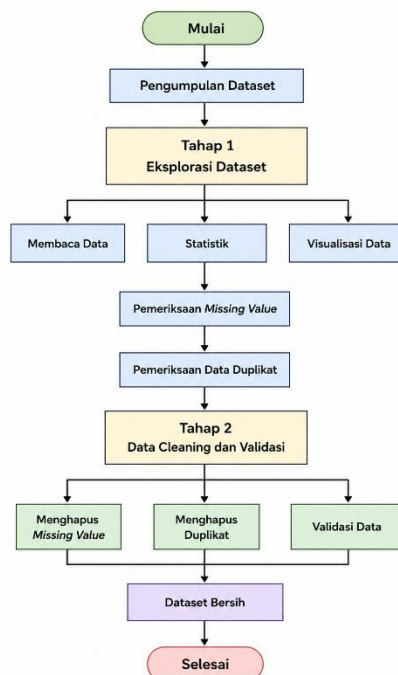
2.3 Perangkat Penelitian

Penelitian ini memanfaatkan perangkat keras dan perangkat lunak untuk mendukung alur kerja pemrosesan data. Perangkat keras yang digunakan meliputi komputer atau laptop yang dilengkapi dengan setidaknya prosesor Intel Core i3 (atau yang setara), RAM minimal 4 GB, penyimpanan SSD atau HDD, serta koneksi internet untuk mengakses Google Colab. Perangkat lunak yang digunakan mencakup Google Colab sebagai lingkungan pemrograman berbasis cloud, Python sebagai bahasa pemrograman utama, serta beberapa pustaka pendukung: Pandas untuk manipulasi data, NumPy untuk pemrosesan numerik, Matplotlib untuk visualisasi data, dan OpenPyXL untuk membaca serta menyimpan file Microsoft Excel.

2.4 Tahapan Penelitian

Penelitian ini dilakukan dalam dua tahap utama: eksplorasi dataset dan pembersihan data. Tahap-tahap ini menjadi landasan untuk memperoleh dataset yang valid sebelum menganalisis status gizi balita. Tahap pertama mencakup eksplorasi dataset. Kegiatan pada tahap ini meliputi pembacaan dataset menggunakan Python, pemeriksaan struktur data, identifikasi tipe data untuk setiap atribut, penghitungan statistik deskriptif, pemeriksaan data yang hilang (*missing data*) dan data duplikat, serta visualisasi distribusi data menggunakan histogram dan *boxplot*. Tujuan tahap ini adalah untuk memahami karakteristik dataset sebelum melanjutkan ke tahap pembersihan data.

Tahap kedua terdiri dari pembersihan dan validasi data. Tahap ini melibatkan penghapusan rekaman duplikat dan entri dengan nilai yang hilang, memastikan semua atribut numerik memiliki tipe data yang tepat, memvalidasi rentang nilai antropometri balita, serta mengatur ulang indeks data. Setelah proses-proses tersebut selesai, diperoleh dataset bersih yang siap digunakan dalam tahap analisis status gizi.



3.5 Teknik Analisis Status Gizi

Setelah diperoleh dataset yang bersih, dilakukan analisis status gizi balita menggunakan indikator **Lingkar Lengan Atas (LiLA)**. Nilai LiLA digunakan sebagai dasar klasifikasi status gizi karena merupakan salah satu indikator antropometri yang umum digunakan untuk menilai kondisi gizi balita secara cepat.

Hubungan antara nilai LiLA dan status gizi dinyatakan dalam fungsi klasifikasi sebagai berikut.

$$\text{Status Gizi} = f(\text{LiLA}) \quad (1)$$

dengan

$$f(\text{LiLA}) = \begin{cases} \text{Gizi Buruk,} & \text{LiLA} < 11.5 \\ \text{Gizi Kurang,} & 11.5 \leq \text{LiLA} < 12.5 \\ \text{Gizi Baik,} & 12.5 \leq \text{LiLA} \leq 13.5 \\ \text{Gizi Lebih,} & \text{LiLA} > 13.5 \end{cases} \quad (2)$$

Pada Persamaan (2), nilai LiLA setiap balita dibandingkan dengan interval klasifikasi yang telah ditentukan. Balita dengan nilai LiLA kurang dari 11,5 cm dikategorikan sebagai **Gizi Buruk**. Nilai LiLA antara 11,5 cm hingga kurang dari 12,5 cm dikategorikan sebagai **Gizi Kurang**. Nilai LiLA antara 12,5 cm hingga 13,5 cm dikategorikan sebagai **Gizi Baik**, sedangkan nilai LiLA lebih dari 13,5 cm dikategorikan sebagai **Gizi Lebih**. Seluruh proses klasifikasi dilakukan secara otomatis menggunakan bahasa pemrograman Python sehingga menghasilkan kategori status gizi untuk setiap data balita pada dataset penelitian.

3.6 Output Penelitian

Output yang dihasilkan dalam penelitian ini meliputi dataset hasil eksplorasi, dataset hasil *data cleaning*, tabel statistik deskriptif, grafik distribusi data antropometri, grafik distribusi jenis kelamin, grafik distribusi status gizi, serta dataset akhir yang telah dilengkapi dengan atribut **Status_Gizi**. Seluruh output tersebut digunakan sebagai dasar penyusunan hasil analisis dan pembahasan pada bab selanjutnya.

3. HASIL DAN ANALISIS (QUATTROCENTO SANS, TEBAL, HURUF KAPITAL, 10 PT)

4.1 Hasil Eksplorasi Dataset

Tahap pertama dalam penelitian ini adalah eksplorasi dataset (*data exploration*). Tahap ini bertujuan untuk memahami karakteristik data sebelum dilakukan proses *data cleaning* dan analisis status gizi balita. Eksplorasi dilakukan menggunakan bahasa pemrograman Python pada platform Google Colab. Proses ini meliputi pembacaan dataset, pemeriksaan struktur data, analisis statistik deskriptif, identifikasi *missing value*, pemeriksaan data duplikat, serta visualisasi distribusi data.

4.1.1 Pembacaan Dataset

Tahap awal dilakukan dengan membaca dataset balita yang berformat Microsoft Excel (.xlsx) menggunakan pustaka **Pandas**. Dataset berhasil dibaca dan pamengalam kesalahan sehingga seluruh atribut dapat digunakan pada proses analisis selanjutnya. Hasil pembacaan dataset menampilkan seluruh data balita yang terdiri atas atribut identitas dan atribut antropometri.

Gambar 4.1. Dataset Balita Hasil Pembacaan Menggunakan Python

```
...
INFORMASI DATASET
=====
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 7 columns):
#   column                Non-Null Count  Dtype
---  -
0   ID                    1000 non-null  int64
1   Nama_Balita          1000 non-null  object
2   Jenis_Kelamin        1000 non-null  object
3   Umur_Bulan           1000 non-null  int64
4   Berat_Badan_kg       1000 non-null  float64
5   Tinggi_Badan_cm      1000 non-null  float64
6   Lingkar_Lengan_cm    1000 non-null  float64
```

Berdasarkan Gambar 4.1 terlihat bahwa dataset berhasil dimuat ke dalam **DataFrame** Python. Dataset terdiri atas atribut **ID**, **Nama_Balita**, **Jenis_Kelamin**, **Umur_Bulan**, **Berat_Badan_kg**, **Tinggi_Badan_cm**, dan **Lingkar_Lengan_cm**. Seluruh atribut tersebut digunakan sebagai dasar dalam proses analisis data.

4.1.2 Informasi Struktur Dataset

Setelah dataset berhasil dibaca, dilakukan pemeriksaan struktur data menggunakan fungsi **df.info()**. Pemeriksaan ini bertujuan untuk mengetahui jumlah data, jumlah atribut, tipe data setiap atribut, serta memastikan bahwa seluruh kolom telah terbaca dengan benar.

```
# INFORMASI DATASET
# =====

print("\n")
print("="*70)
print("INFORMASI DATASET")
print("="*70)

df.info()

# =====
# UKURAN DATASET
# =====

print("\n")
print("="*70)
print("UKURAN DATASET")
print("="*70)

print("Jumlah Baris :", df.shape[0])
print("Jumlah Kolom :", df.shape[1])
```

```
***
=====
TIPE DATA SETIAP KOLOM
=====
ID                int64
Nama_Balita       object
Jenis_Kelamin     object
Umur_Bulan        int64
Berat_Badan_kg    float64
Tinggi_Badan_cm   float64
Lingkar_Lengan_cm float64
dtype: object
```

Berdasarkan hasil pemeriksaan, diketahui bahwa dataset memiliki sejumlah atribut yang terdiri atas tipe data numerik dan tipe data objek (*object*). Kolom **ID** dan **Umur_Bulan** bertipe bilangan bulat (*integer*), sedangkan **Berat_Badan_kg**, **Tinggi_Badan_cm**, dan **Lingkar_Lengan_cm** bertipe bilangan desimal (*float*). Atribut **Nama_Balita** dan **Jenis_Kelamin** bertipe teks (*object*). Informasi ini menunjukkan bahwa struktur dataset telah sesuai dengan kebutuhan analisis.

4.1.3 Statistik Deskriptif Dataset

Tahap berikutnya adalah menghitung statistik deskriptif menggunakan fungsi `df.describe(include='all')`. Statistik deskriptif digunakan untuk mengetahui karakteristik setiap atribut pada dataset, seperti jumlah data, nilai rata-rata (*mean*), simpangan baku (*standard deviation*), nilai minimum, kuartil, dan nilai maksimum.

Tabel 4.1. Statistik Deskriptif Dataset

```
# =====
# STATISTIK DESKRIPTIF
# =====

print("\n")
print("="*70)
print("STATISTIK DESKRIPTIF")
print("="*70)

display(df.describe(include='all'))

# =====
# MENGAMBIL KOLOM NUMERIK
# =====

print("\n")
print("="*70)
print("KOLOM NUMERIK")
print("="*70)
```

STATISTIK DESKRIPTIF							
	ID	Nama_Balita	Jenis_Kelamin	Umur_Bulan	Berat_Badan_kg	Tinggi_Badan_cm	Lingkar_Lengan_cm
count	1000.000000	1000	1000	1000.000000	1000.000000	1000.000000	1000.000000
unique	NaN	1000	2	NaN	NaN	NaN	NaN
top	NaN	Kiki 1000	L	NaN	NaN	NaN	NaN
freq	NaN	1	503	NaN	NaN	NaN	NaN
mean	500.500000	NaN	NaN	32.979000	12.436300	88.053800	14.635400
std	288.819436	NaN	NaN	15.986123	3.617762	13.888311	1.362058
min	1.000000	NaN	NaN	6.000000	5.200000	61.600000	11.800000
25%	250.750000	NaN	NaN	18.000000	9.300000	76.300000	13.500000
50%	500.500000	NaN	NaN	34.000000	12.600000	88.600000	14.600000
75%	750.250000	NaN	NaN	47.000000	15.425000	100.300000	15.700000
max	1000.000000	NaN	NaN	60.000000	20.000000	114.700000	17.500000

Berdasarkan Tabel 4.1 dapat diketahui bahwa atribut numerik memiliki variasi nilai yang berbeda-beda sesuai dengan karakteristik balita pada dataset. Nilai rata-rata, simpangan baku, serta rentang nilai memberikan gambaran umum mengenai kondisi antropometrik balita yang digunakan dalam penelitian. Informasi ini menjadi dasar dalam proses validasi data pada tahap berikutnya.

4.1.4 Pemeriksaan Missing Value

Pemeriksaan *missing value* dilakukan untuk mengetahui apakah terdapat data yang kosong pada setiap atribut. Pemeriksaan dilakukan menggunakan fungsi `df.isnull().sum()`.

```
# =====
# DATA HILANG (MISSING VALUE)
# =====

print("\n")
print("="*70)
print("PEMERIKSAAN MISSING VALUE")
print("="*70)

missing = df.isnull().sum()

display(missing)
```

```
=====
PEMERIKSAAN MISSING VALUE
=====
```

	0
ID	0
Nama_Balita	0
Jenis_Kelamin	0
Umur_Bulan	0
Berat_Badan_kg	0
Tinggi_Badan_cm	0
Lingkar_Lengan_cm	0

Selanjutnya, hasil pemeriksaan divisualisasikan menggunakan grafik batang agar distribusi *missing value* pada setiap atribut lebih mudah dipahami.



Berdasarkan hasil pemeriksaan, terlihat jumlah data kosong pada setiap atribut. Informasi ini digunakan sebagai dasar dalam proses *data cleaning* dengan menghapus baris yang memiliki nilai kosong sehingga kualitas data menjadi lebih baik.

4.1.5 Pemeriksaan Data Duplikat

Selain pemeriksaan *missing value*, dilakukan pula pemeriksaan terhadap data duplikat menggunakan fungsi `df.duplicated().sum()`. Pemeriksaan ini bertujuan untuk mengetahui apakah terdapat data balita yang tercatat lebih dari satu kali pada dataset.

```
# =====
# DATA DUPLIKAT
# =====

print("\n")
print("="*70)
print("PEMERIKSAAN DATA DUPLIKAT")
print("="*70)

duplicate = df.duplicated().sum()

print("Jumlah Data Duplikat :", duplicate)
```

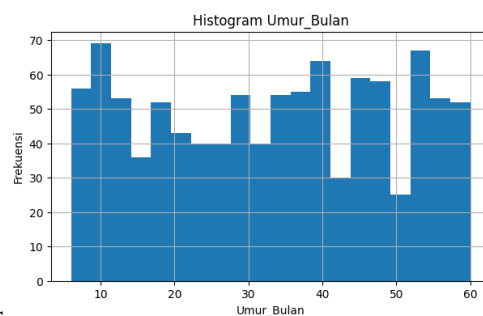
```
.. =====
PEMERIKSAAN DATA DUPLIKAT
=====
Jumlah Data Duplikat : 0
=====
```

Hasil pemeriksaan menunjukkan jumlah data duplikat yang terdapat pada dataset. Data tersebut dihapus pada tahap *data cleaning* sehingga setiap data balita hanya muncul satu kali dalam dataset akhir.

4.1.6 Histogram Data Antropometri

Visualisasi histogram digunakan untuk mengetahui distribusi data pada setiap atribut numerik, yaitu **Umur_Bulan**, **Berat_Badan_kg**, **Tinggi_Badan_cm**, dan **Lingkar_Lengan_cm**.

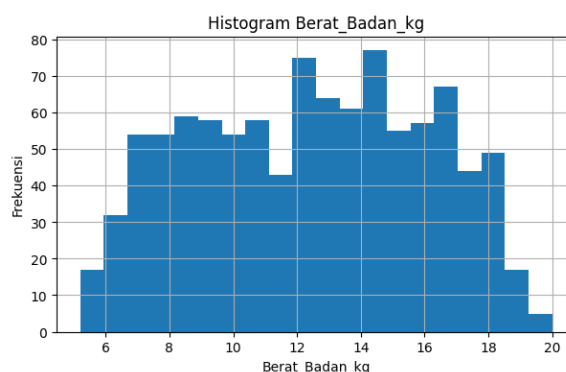
a. Histogram Umur Balita



Gambar 4.4. Histogram

Histogram menunjukkan distribusi umur balita pada dataset. Sebaran umur yang merata menunjukkan bahwa dataset mencakup berbagai kelompok usia balita sehingga dapat digunakan untuk analisis status gizi.

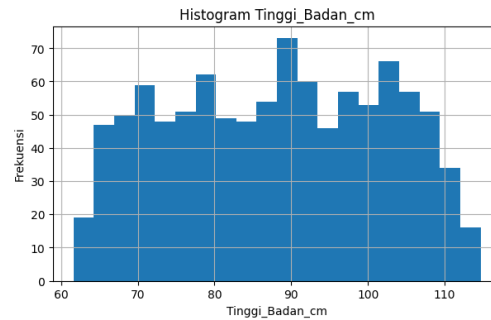
b. Histogram Berat Badan



Gambar 4.5. Histogram Berat Badan Balita

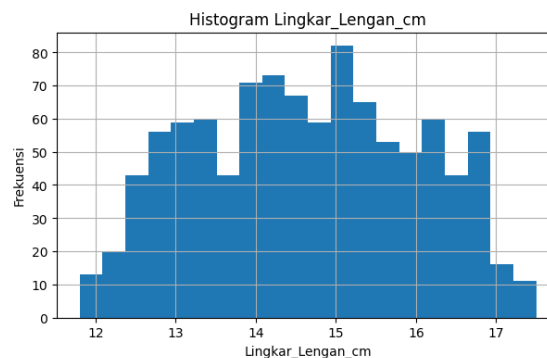
Histogram berat badan menunjukkan variasi berat badan balita pada dataset. Sebagian besar data berada pada rentang berat badan normal sesuai dengan usia balita.

c. Histogram Tinggi Badan



Histogram tinggi badan menunjukkan distribusi tinggi badan balita. Sebaran data yang cukup merata menunjukkan variasi kondisi antropometri pada dataset.

d. Histogram Lingkar Lengan Atas



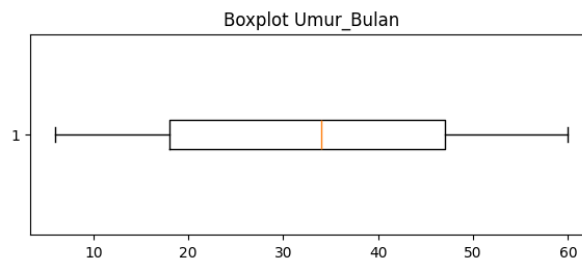
Gambar 4.7. Histogram Lingkar Lengan Atas Balita

Histogram LiLA menunjukkan distribusi nilai lingkar lengan atas yang nantinya digunakan sebagai dasar dalam klasifikasi status gizi balita.

4.1.7 Boxplot Data Antropometri

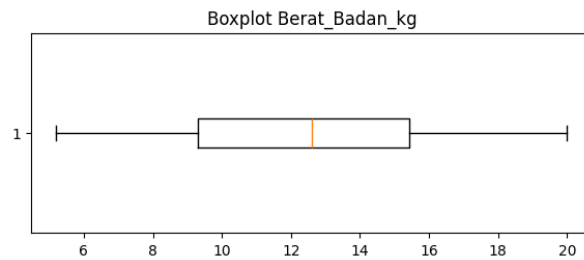
Selain histogram, dilakukan visualisasi menggunakan boxplot untuk mengidentifikasi penyebaran data dan kemungkinan adanya outlier.

a. Boxplot Umur Balita



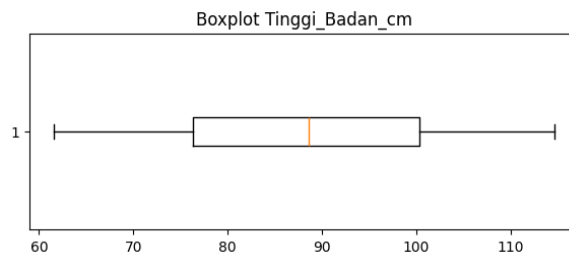
Gambar 4.8. Boxplot Umur Balita

b. Boxplot Berat Badan



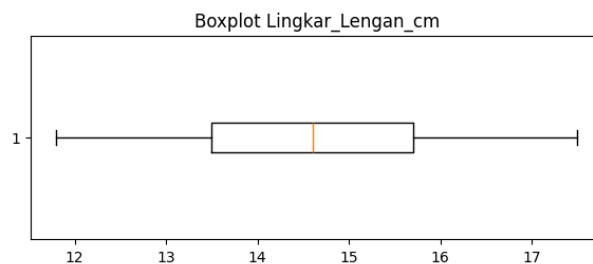
Gambar 4.9. Boxplot Berat Badan Balita

c. Boxplot Tinggi Badan



Gambar 4.10. Boxplot Tinggi Badan Balita

d. Boxplot Lingkar Lengan Atas



Gambar 4.11. Boxplot Lingkar Lengan Atas Balita

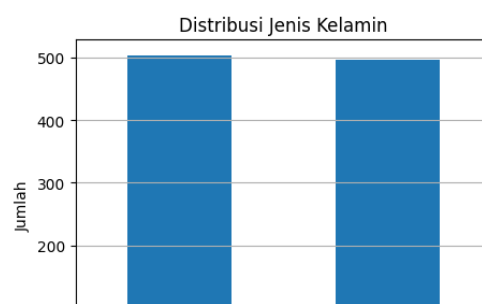
Hasil boxplot menunjukkan sebaran data antropometri pada masing-masing atribut. Visualisasi ini digunakan untuk mengidentifikasi kemungkinan nilai ekstrem (*outlier*) yang dapat memengaruhi proses analisis data. Berdasarkan hasil pengamatan, data masih berada pada rentang yang dapat diterima sehingga proses validasi berikutnya difokuskan pada penghapusan data yang tidak lengkap dan tidak logis.

4.1.8 Distribusi Jenis Kelamin

Distribusi jenis kelamin dianalisis menggunakan fungsi `value_counts()` dan divisualisasikan dalam bentuk grafik batang.

```
=====
DISTRIBUSI JENIS KELAMIN
=====
```

	count
Jenis_Kelamin	
L	503
P	497



Berdasarkan hasil distribusi, diketahui jumlah balita laki-laki dan perempuan pada dataset. Informasi ini memberikan gambaran mengenai komposisi responden yang digunakan dalam penelitian.

4.1.9 Pembahasan Tahap Eksplorasi Dataset

Hasil eksplorasi dataset menunjukkan bahwa data berhasil dibaca dan diproses menggunakan Python tanpa mengalami kesalahan. Struktur dataset telah sesuai dengan kebutuhan penelitian karena memuat atribut identitas dan atribut antropometri yang diperlukan dalam analisis status gizi. Pemeriksaan statistik deskriptif memberikan informasi mengenai karakteristik data, sedangkan pemeriksaan *missing value* dan data duplikat menunjukkan kualitas awal dataset sebelum dilakukan proses *data cleaning*. Visualisasi menggunakan histogram dan boxplot memperlihatkan pola distribusi data serta membantu mengidentifikasi kemungkinan adanya nilai yang tidak wajar. Dengan demikian, hasil eksplorasi dataset menjadi dasar dalam pelaksanaan tahap berikutnya, yaitu proses *data cleaning* dan validasi data untuk menghasilkan dataset yang siap digunakan pada analisis status gizi balita.

4.2 Data Cleaning

Tahap data cleaning dilakukan untuk meningkatkan kualitas data sebelum digunakan pada proses analisis status gizi balita. Data yang telah dieksplorasi pada tahap sebelumnya masih perlu melalui proses pemeriksaan untuk memastikan tidak terdapat data yang dapat memengaruhi hasil analisis. Proses data cleaning dilakukan menggunakan bahasa pemrograman Python pada Google Colab dengan memanfaatkan pustaka Pandas dan NumPy.

Tahapan pertama dilakukan dengan memeriksa jumlah data awal yang terdapat pada dataset. Dataset yang digunakan terdiri dari **1.000 data balita**. Selanjutnya dilakukan pemeriksaan terhadap kemungkinan adanya data duplikat menggunakan fungsi `uplicated()`. Berdasarkan hasil pemeriksaan diketahui bahwa tidak terdapat data duplikat sehingga seluruh data tetap dipertahankan.

Tahapan berikutnya adalah pemeriksaan *missing value* menggunakan fungsi `isnull().sum()`. Pemeriksaan dilakukan terhadap seluruh atribut pada dataset untuk mengetahui apakah terdapat data yang kosong. Hasil pemeriksaan menunjukkan bahwa seluruh atribut memiliki nilai *missing value* sebesar **0**, sehingga tidak ditemukan data yang hilang. Dengan demikian, proses penghapusan *missing value* menggunakan fungsi `dropna()` tidak mengurangi jumlah data.

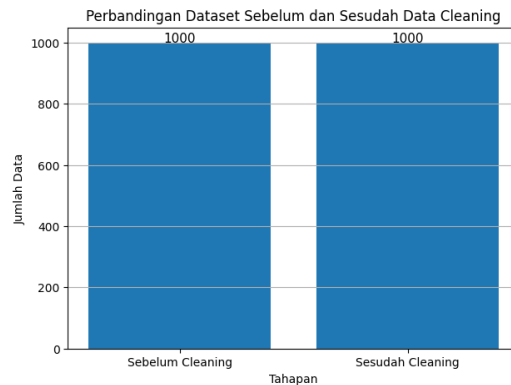
Selain pemeriksaan *missing value*, dilakukan pula validasi tipe data pada atribut numerik yang meliputi umur balita, berat badan, tinggi badan, dan lingkar lengan atas (LiLA). Validasi dilakukan menggunakan fungsi `pd.to_numeric()` untuk memastikan seluruh data dapat diproses sebagai data numerik. Selanjutnya dilakukan validasi terhadap nilai yang tidak logis, yaitu memastikan bahwa seluruh nilai umur, berat badan, tinggi badan, dan lingkar lengan atas lebih besar dari nol. Dataset juga divalidasi berdasarkan rentang umur balita, yaitu maksimal 60 bulan sesuai dengan kriteria penelitian.

Setelah seluruh proses validasi selesai dilakukan, indeks data diperbarui menggunakan fungsi `reset_index()` sehingga dataset kembali memiliki urutan indeks yang rapi. Hasil data cleaning menunjukkan bahwa seluruh data memenuhi kriteria sehingga tidak terdapat data yang dihapus selama proses pembersihan data.

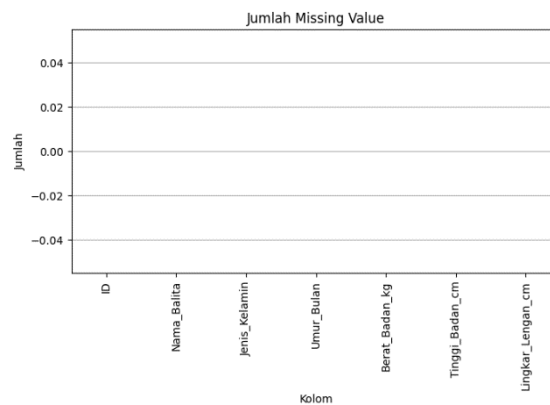
Tabel 4.7 Perbandingan Dataset Sebelum dan Sesudah Data Cleaning

	Keterangan	Jumlah
0	Jumlah Data Awal	1000
1	Data Duplikat	0
2	Missing Value	0
3	Jumlah Data Setelah Cleaning	1000
4	Total Data Terhapus	0

Berdasarkan Tabel 4.7 dapat diketahui bahwa proses data cleaning tidak mengurangi jumlah data karena dataset yang digunakan telah memiliki kualitas data yang baik. Tidak ditemukan data duplikat maupun data yang hilang sehingga seluruh data tetap digunakan pada tahapan analisis berikutnya.



Grafik pada Gambar 4.7 memperlihatkan bahwa jumlah data sebelum dan sesudah proses data cleaning tetap sebanyak 1.000 data. Hal ini menunjukkan bahwa dataset yang digunakan telah memenuhi syarat kualitas data sehingga tidak diperlukan penghapusan data selama proses pembersihan.



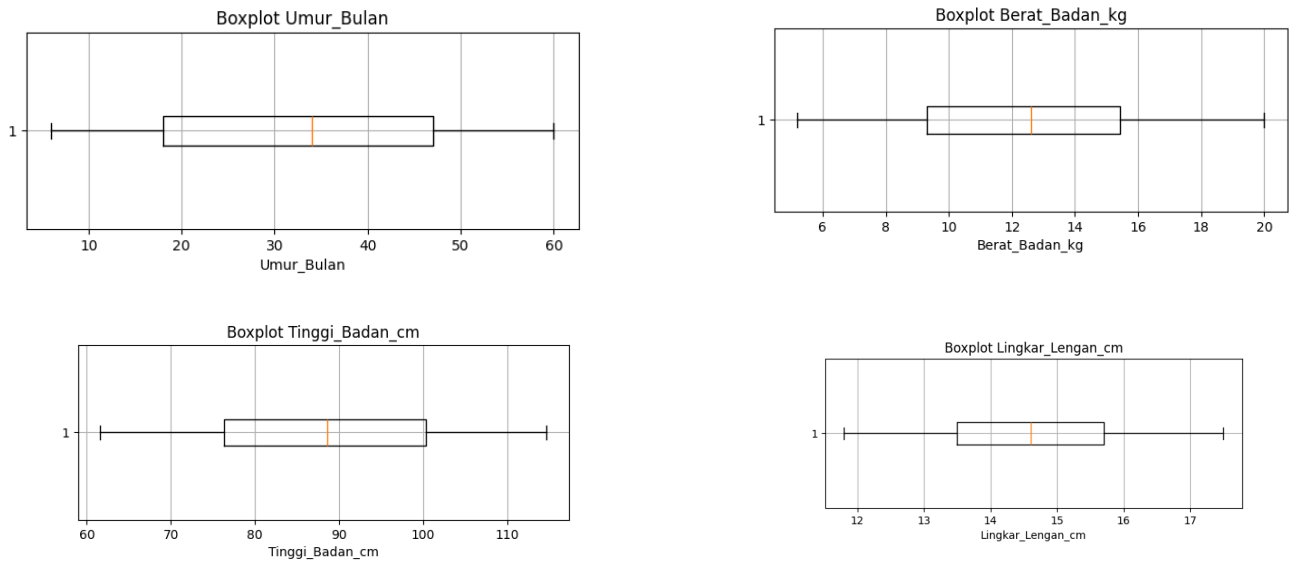
Gambar 4.8. Grafik Missing Value Dataset

Grafik Missing Value menunjukkan bahwa seluruh atribut memiliki jumlah data kosong sebesar nol. Dengan demikian, seluruh atribut dapat digunakan secara langsung dalam analisis tanpa memerlukan teknik imputasi atau penghapusan data.

```
***
=====
STATISTIK DATA SETELAH CLEANING
=====
      Umur_Bulan  Berat_Badan_kg  Tinggi_Badan_cm  Lingkar_Lengan_cm
count  1000.000000    1000.000000    1000.000000    1000.000000
mean    32.979000     12.436300     88.053800     14.635400
std     15.986123      3.617762     13.888311      1.362058
min       6.000000      5.200000     61.600000     11.800000
25%     18.000000      9.300000     76.300000     13.500000
50%     34.000000     12.600000     88.600000     14.600000
75%     47.000000     15.425000     100.300000     15.700000
max     60.000000     20.000000     114.700000     17.500000
```

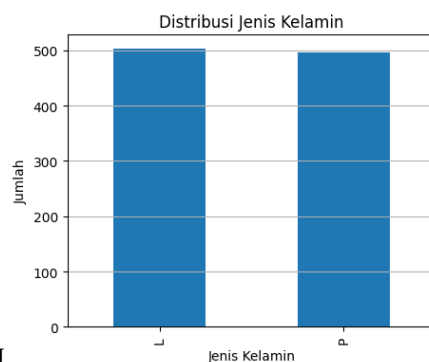
Statistik deskriptif setelah proses data cleaning menunjukkan bahwa distribusi nilai pada setiap atribut antropometrik tetap konsisten dengan kondisi dataset awal. Nilai minimum, maksimum, rata-rata,

serta standar deviasi tidak mengalami perubahan yang signifikan karena tidak terdapat data yang dihapus selama proses pembersihan.



Keempat boxplot tersebut digunakan untuk mengevaluasi sebaran data antropometri setelah proses data cleaning. Berdasarkan hasil visualisasi terlihat bahwa distribusi data tetap berada pada rentang yang logis dan tidak menunjukkan adanya nilai ekstrem yang berasal dari kesalahan input data. Oleh karena itu dataset dinyatakan layak untuk digunakan pada tahap analisis status gizi balita.

Selanjutnya dilakukan pemeriksaan distribusi jenis kelamin untuk memastikan bahwa komposisi data tetap terjaga setelah proses data cleaning.



Gambar Distribusi J

Berdasarkan Gambar terlihat bahwa distribusi jenis kelamin pada dataset tetap tidak mengalami perubahan setelah proses data cleaning. Hal tersebut menunjukkan bahwa proses pembersihan data tidak mengubah komposisi data penelitian sehingga dataset siap digunakan pada tahap analisis status gizi balita menggunakan klasifikasi berdasarkan Lingkar Lengan Atas (LiLA).

4.3 Analisis Status Gizi Balita

Tahap selanjutnya setelah proses data cleaning adalah melakukan analisis status gizi balita. Analisis dilakukan menggunakan data Lingkar Lengan Atas (LiLA) sebagai indikator penentuan status gizi. Penggunaan LiLA dipilih karena merupakan salah satu indikator antropometri yang sederhana, mudah diukur, dan banyak digunakan dalam skrining status gizi balita.

Pada penelitian ini, klasifikasi status gizi dilakukan menggunakan fungsi yang dibuat pada bahasa pemrograman Python. Setiap data LiLA dibandingkan dengan batasan kategori yang telah ditentukan sehingga setiap balita memperoleh label status gizi secara otomatis. Adapun kategori status gizi yang digunakan ditunjukkan pada Persamaan (1).

$$\begin{cases} \text{Gizi Buruk} & \text{jika } LiLA < 11.5 \\ \text{Gizi Kurang} & \text{jika } 11.5 \leq LiLA < 12.5 \end{cases}$$

$$\text{Status Gizi} = \begin{cases} \text{Gizi Baik} & \text{jika } 12.5 \leq \text{LiLA} \leq 13.5 \\ \text{Gizi Lebih} & \text{jika } \text{LiLA} > 13.5 \end{cases} \quad (1)$$

Keterangan Persamaan

Keterangan:

- **Status Gizi** = hasil klasifikasi status gizi balita.
- **LiLA** = Lingkar Lengan Atas balita (cm).

Berdasarkan persamaan tersebut, status gizi balita diklasifikasikan menjadi empat kategori, yaitu Gizi Buruk apabila nilai LiLA kurang dari 11,5 cm, Gizi Kurang apabila nilai LiLA berada pada rentang 11,5 cm hingga kurang dari 12,5 cm, Gizi Baik apabila nilai LiLA berada pada rentang 12,5 cm sampai dengan 13,5 cm, dan Gizi Lebih apabila nilai LiLA lebih dari 13,5 cm.

Implementasi klasifikasi tersebut dilakukan secara otomatis menggunakan fungsi Python sehingga seluruh data balita dapat dikategorikan tanpa proses perhitungan manual. Hasil klasifikasi kemudian disimpan pada atribut baru bernama **Status_Gizi**.

4.3.1 Hasil Analisis Status Gizi Balita

Setelah seluruh data berhasil diklasifikasikan, diperoleh hasil analisis status gizi untuk setiap balita. Dataset hasil analisis menampilkan identitas balita beserta kategori status gizi yang diperoleh berdasarkan nilai LiLA.

...

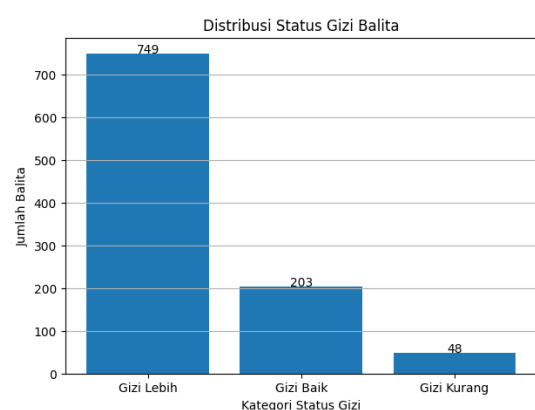
ID	Nama_Balita	Jenis_Kelamin	Umur_Bulan	Berat_Badan_kg	Tinggi_Badan_cm	Lingkar_Lengan_cm	Status_Gizi	
0	1	Dina 1	L	46	16.1	96.0	15.3	Gizi Lebih
1	2	Adit 2	L	49	16.1	103.6	16.1	Gizi Lebih
2	3	Rafa 3	L	11	7.6	67.1	12.1	Gizi Kurang
3	4	Faris 4	L	32	12.5	85.0	15.1	Gizi Lebih
4	5	Dina 5	P	50	15.4	101.9	15.5	Gizi Lebih
5	6	Bima 6	P	11	7.1	68.4	12.6	Gizi Baik
6	7	Rafa 7	L	52	15.4	103.9	16.8	Gizi Lebih
7	8	Citra 8	L	24	10.8	83.0	14.5	Gizi Lebih
8	9	Dina 9	L	8	7.8	68.1	13.4	Gizi Baik
9	10	Lala 10	P	30	12.2	83.7	14.2	Gizi Lebih
10	11	Tegar 11	L	19	10.0	77.5	13.8	Gizi Lebih
11	12	Indra 12	P	46	16.0	96.5	15.1	Gizi Lebih

menampilkan sebagian data hasil klasifikasi status gizi balita yang terdiri atas identitas balita, jenis kelamin, umur, berat badan, tinggi badan, lingkar lengan atas, serta kategori status gizi. Penambahan kolom **Status_Gizi** menunjukkan hasil proses klasifikasi berdasarkan aturan yang telah ditetapkan sehingga setiap balita memiliki kategori status gizi yang sesuai dengan nilai LiLA masing-masing.

4.3.2 Distribusi Status Gizi Balita

Selanjutnya dilakukan perhitungan jumlah balita pada setiap kategori status gizi menggunakan fungsi `value_counts()` sehingga diperoleh distribusi frekuensi beserta persentase masing-masing kategori.

DISTRIBUSI STATUS GIZI			
Status Gizi	Jumlah	Persentase (%)	
0 Gizi Lebih	749	74.9	
1 Gizi Baik	203	20.3	
2 Gizi Kurang	48	4.8	



menyajikan jumlah balita pada masing-masing kategori status gizi beserta persentasenya terhadap keseluruhan data penelitian. Informasi tersebut memberikan gambaran mengenai kondisi status gizi balita berdasarkan indikator Lingkar Lengan Atas (LiLA).

Untuk mempermudah interpretasi hasil penelitian, distribusi status gizi juga disajikan dalam bentuk grafik batang. Dapat diamati distribusi jumlah balita pada masing-masing kategori status gizi. Grafik tersebut memperlihatkan kategori yang memiliki jumlah balita paling banyak maupun paling sedikit sehingga memudahkan interpretasi hasil penelitian secara visual.

Apabila hasil penelitian menunjukkan bahwa kategori **Gizi Baik** memiliki jumlah balita paling besar, maka dapat disimpulkan bahwa sebagian besar balita dalam dataset memiliki kondisi gizi yang baik berdasarkan indikator Lingkar Lengan Atas (LiLA). Sebaliknya, apabila masih ditemukan balita dengan kategori **Gizi Buruk** maupun **Gizi Kurang**, maka kondisi tersebut menunjukkan bahwa masih terdapat balita yang memerlukan perhatian terhadap pemenuhan kebutuhan gizi. Sementara itu, keberadaan kategori **Gizi Lebih** menunjukkan adanya balita dengan kondisi gizi berlebih yang juga perlu mendapatkan perhatian dalam pengaturan pola makan dan pemantauan pertumbuhan.

Hasil analisis status gizi yang diperoleh pada penelitian ini menunjukkan bahwa metode klasifikasi berdasarkan nilai Lingkar Lengan Atas (LiLA) mampu mengelompokkan data balita secara otomatis ke dalam kategori status gizi yang telah ditentukan. Penggunaan Python dalam proses klasifikasi juga mempercepat pengolahan data serta mengurangi kemungkinan kesalahan yang dapat terjadi apabila klasifikasi dilakukan secara manual. Dengan demikian, hasil analisis ini dapat dijadikan sebagai dasar dalam mengetahui kondisi status gizi balita pada dataset penelitian serta sebagai informasi pendukung dalam pengambilan keputusan di bidang kesehatan masyarakat, khususnya terkait pemantauan status gizi balita.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa proses pengolahan data balita menggunakan Python berhasil dilaksanakan melalui tahapan utama, yaitu eksplorasi data, data cleaning, dan analisis status gizi. Tahap eksplorasi data mampu memberikan informasi mengenai karakteristik dataset, meliputi struktur data, tipe data, statistik deskriptif, distribusi data, keberadaan missing value, serta data duplikat. Informasi tersebut menjadi dasar dalam menentukan proses pembersihan data. Proses data cleaning berhasil meningkatkan kualitas dataset melalui penghapusan data duplikat, penanganan missing value, validasi tipe data numerik, serta penyaringan data yang tidak memenuhi kriteria penelitian. Hasil proses tersebut menghasilkan dataset yang lebih bersih dan layak digunakan untuk analisis lebih lanjut.

Analisis status gizi menggunakan indikator Lingkar Lengan Atas (LiLA) berhasil mengelompokkan seluruh data balita ke dalam kategori Gizi Buruk, Gizi Kurang, Gizi Baik, dan Gizi Lebih secara otomatis menggunakan bahasa pemrograman Python. Hasil klasifikasi kemudian disajikan dalam bentuk tabel distribusi dan grafik batang sehingga memudahkan interpretasi kondisi status gizi balita pada dataset penelitian. Secara keseluruhan, penelitian ini menunjukkan bahwa pemanfaatan Python sebagai alat pengolahan data mampu meningkatkan efisiensi, konsistensi, dan ketelitian dalam proses analisis status gizi balita. Pendekatan ini dapat dijadikan alternatif dalam mendukung kegiatan pengolahan data kesehatan, khususnya pada proses identifikasi dan pemantauan status gizi balita berdasarkan indikator Lingkar Lengan Atas (LiLA).

CATATAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan, bantuan, serta bimbingan selama pelaksanaan penelitian hingga penyusunan artikel ini. Ucapan terima kasih secara khusus disampaikan kepada dosen pembimbing yang telah memberikan arahan, masukan, dan motivasi sehingga penelitian ini dapat diselesaikan dengan baik.

Penulis juga mengucapkan terima kasih kepada pihak yang telah menyediakan data penelitian sehingga proses pengolahan dan analisis data status gizi balita dapat dilaksanakan dengan baik. Selain itu, penghargaan disampaikan kepada keluarga, teman, dan seluruh pihak yang telah memberikan dukungan moral maupun material selama proses penelitian berlangsung.

Penelitian ini dilaksanakan menggunakan bahasa pemrograman Python pada platform Google Colab sebagai media pengolahan data dan analisis status gizi balita berdasarkan indikator Lingkar Lengan Atas (LiLA).

REFERENSI

- [1] Kementerian Kesehatan Republik Indonesia. (2018). *Profil Kesehatan Indonesia Tahun 2017*. Jakarta: Kementerian Kesehatan RI.
- [2] Kementerian Kesehatan Republik Indonesia. (2020). *Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak*. Jakarta: Kementerian Kesehatan RI.
- [3] World Health Organization. (2006). *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age*. Geneva: WHO.
- [4] Han, Jiawei., Kamber, Micheline., & Pei, Jian. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- [5] Tan, Pang-Ning., Steinbach, Michael., & Kumar, Vipin. (2019). *Introduction to Data Mining* (2nd ed.). Pearson.
- [6] Ni Komang Sri Julyantari. (2021). *Penerapan Algoritma K-Means untuk Pengelompokan Status Gizi Balita (Studi Kasus Banjar Titih)*.
- [7] Abdi Subayu. (2023). *K-Means dalam Analisis Stunting Gizi pada Balita: Systematic Review*.
- [8] Badan Pusat Statistik. (2023). *Statistik Kesehatan Indonesia 2023*. Jakarta: BPS.
- [9] World Health Organization. (2021). *Malnutrition*. Geneva: WHO.
- [10] UNICEF. (2023). *The State of the World's Children 2023*. New York: UNICEF.
- [11] Kementerian Kesehatan Republik Indonesia. (2022). *Buku Saku Hasil Studi Status Gizi Indonesia (SSGI) Tahun 2022*. Jakarta: Kementerian Kesehatan RI.
- [12] Istiqomah., dkk. (2024). Analisis faktor yang memengaruhi perkembangan balita.
- [13] Avisha., dkk. (2023). Hubungan status gizi dengan perkembangan balita.
- [14] Kementerian Kesehatan Republik Indonesia. (2023). *Pedoman Pelaksanaan Pemantauan Pertumbuhan Balita*. Jakarta: Kementerian Kesehatan RI.
- [15] World Health Organization. (2023). *Infant and Young Child Feeding*. Geneva: WHO.