Penggunaan Metode Bootstrap dalam Analisis Perbandingan Statistik Deskriptif Indeks Pembangunan Literasi Masyarakat (IPLM) di Wilayah Sumatera Tahun 2024

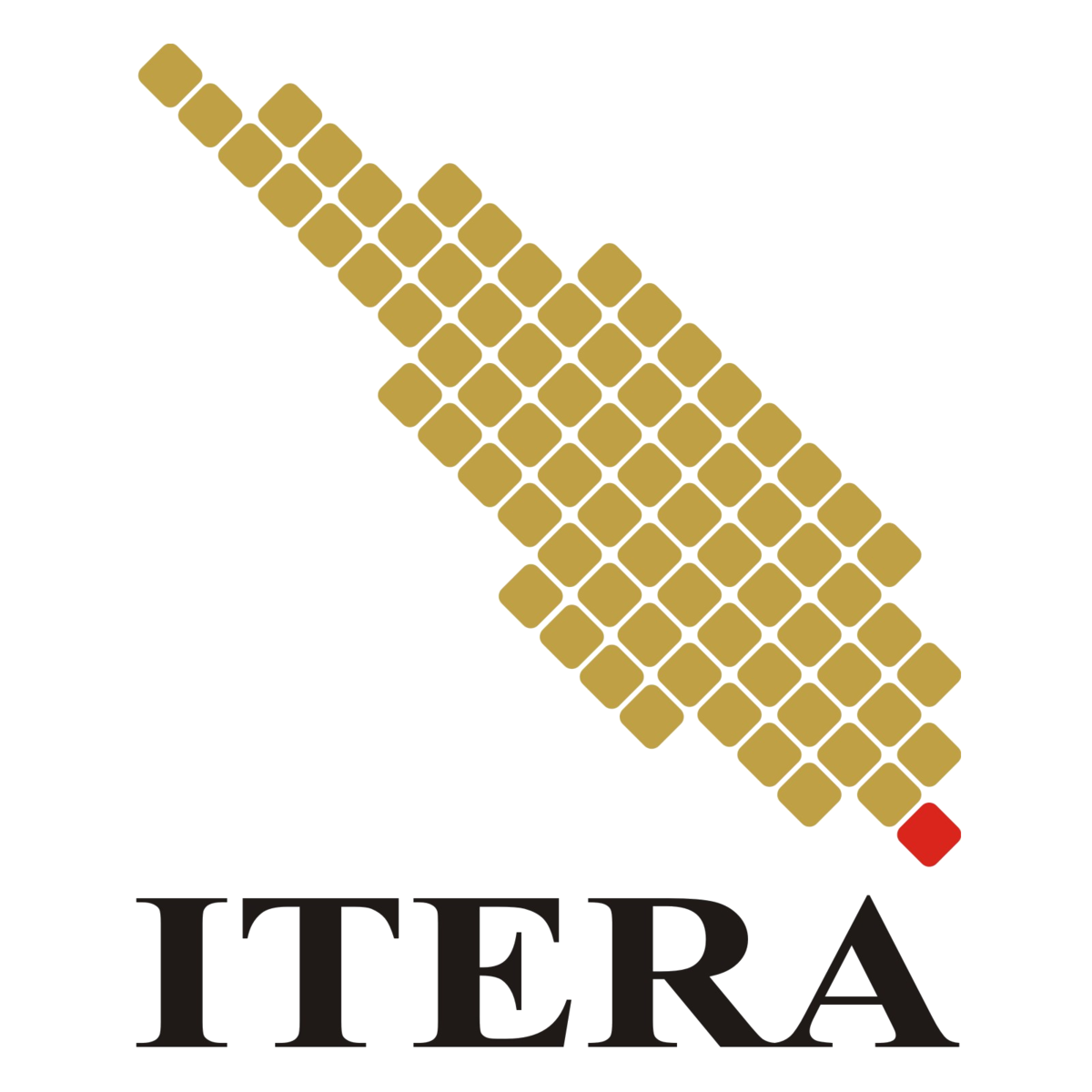
**Tugas Besar**

**Komputasi Statistik**

Kelompok Ke - 1

Nama Anggota Kelompok:

1. Nabyla Sharfina (123450008)
2. Hafsa Fazila Arradhi (123450079)
3. Fabio Banyu Cyto (123450104)
4. Muhammad Naufal Ramadhan (123450113)



**PROGRAM STUDI SAINS DATA**

**FAKULTAS SAINS**

**INSTITUT TEKNOLOGI SUMATERA**

**2025**

# DAFTAR ISI

[**DAFTAR ISI ii**](#_k0o99vxstf0j)

[**ABSTRAK 1**](#_k0o99vxstf0j)

[**BAB I 2**](#_gq6xpb1sxjlq)

[1.1. Latar Belakang 2](#_s84iasapyp6i)

[1.2. Rumusan Masalah 3](#_doaqel4mix6f)

[1.3. Tujuan Penelitian 3](#_8b53x0ci8lt4)

[1.4. Manfaat Penelitian 3](#_npaxqilr1i7k)

[1.4.1. Manfaat Praktis: 4](#_snbh2esebud3)

[1.4.2. Manfaat Akademis: 4](#_m1c52wfcnl1s)

[**BAB II 5**](#_cr4zmxkiif01)

[2.1. Teknik Resampling 5](#_idq3hvpmeb37)

[2.2. Bootstrap 7](#_bsufvs7ow5ax)

[2.3. Penilaian Berdasarkan Statistika Deskriptif: Mean, Median, Standar Deviasi 8](#_u4u7ui5vizy)

[2.3.1. Mean 9](#_eb7p4324yopd)

[2.3.2. Median 9](#_v0wss9snuf8c)

[2.3.3. Standar Deviasi 9](#_1a0lcpxebrgy)

[2.4. Indeks Pembangunan Literasi Masyarakat 10](#_f357atcqlhrp)

[2.4.1. Perkembangan IPLM di Indonesia 10](#_at92uyfvx3mf)

[2.4.2. IPLM di Pulau Sumatera 11](#_dshnpq5hlcp7)

[**BAB III 12**](#_xzflqyryfds7)

[3.1. Jenis Penelitian 12](#_7rcnlt8qjck1)

[3.2. Jenis Data 12](#_2egdjtjo4ujo)

[3.3. Teknik Pengumpulan Data 12](#_g0ampascq1x0)

[3.4. Variabel yang Diamati 12](#_1c3pknc5e1sk)

[3.5. Diagram Alir 15](#_59pkvukkamlk)

[3.6. Tools dan Software 16](#_3gbbqgpcvvi4)

**BAB IV 17**

[4.1. Statistik Deskriptif Sebelum Bootstrap 17](#_84n505ym06et)

[4.2. Proses Resampling Bootstrap 19](#_uecw49c4ch2u)

[4.3. Statistik Deskriptif Setelah Bootstrap 20](#_wpo23cmt24r)

[4.4. Interval Kepercayaan Bootstrap 21](#_fcrfjcfuwhu)

[4.5. Visualisasi Hasil Bootstrap 22](#_axzfyttrxz9e)

[4.6. Perbandingan Statistik Sebelum dan Setelah Bootstrap 24](#_mvy94twuwrc2)

[4.7. Diskusi Temuan Penelitian 25](#_fslaygqnp9r7)

**BAB V 26**

[5.1. Kesimpulan 26](#_9awh9va48u2d)

[5.2. Saran 26](#_7rtigid3hadu)

**DAFTAR PUSTAKA 28**

**ABSTRAK**

Analisis terhadap Indeks Pembangunan Literasi Masyarakat (IPLM) tahun 2024 pada wilayah kabupaten dan kota di Pulau Sumatera dilakukan melalui pemaduan statistik deskriptif dan metode bootstrap untuk memperoleh pemahaman yang lebih komprehensif mengenai kondisi literasi regional. Tahap awal analisis mencakup proses pembersihan data, pemeriksaan konsistensi, serta identifikasi nilai ekstrim yang berpotensi mempengaruhi hasil perhitungan. Setelah data siap diolah, dilakukan perhitungan statistik dasar seperti *mean*, *median*, standar deviasi, serta sebaran nilai untuk menggambarkan variasi tingkat literasi antarwilayah. Penggunaan metode bootstrap sebagai teknik resampling nonparametrik memberikan nilai tambah penting, terutama karena mampu menghasilkan distribusi estimasi *mean* dan *median* yang lebih stabil meskipun data memiliki ukuran sampel yang tidak terlalu besar atau menunjukkan indikasi ketidaknormalan. Hasil resampling memperlihatkan bahwa estimasi parameter cenderung mendekati distribusi normal dan memiliki interval kepercayaan yang relatif sempit, yang berarti ketidakpastian estimasi cukup rendah. Konsistensi antara nilai sebelum dan sesudah bootstrap, terutama pada *mean* dan *median*, menjadi indikator bahwa kondisi variabilitas literasi di Sumatera relatif seragam dan tidak menunjukkan perbedaan ekstrem antarwilayah. Sementara itu, perubahan pada standar deviasi masih tetap berada dalam rentang yang wajar sehingga tidak mengubah pola umum persebaran. Temuan ini menegaskan bahwa metode bootstrap mampu memperkuat keandalan hasil analisis dan memberikan landasan yang lebih solid bagi proses interpretasi. Secara keseluruhan, hasil penelitian dapat dimanfaatkan sebagai acuan statistik bagi perumusan strategi dan kebijakan peningkatan literasi di tingkat daerah maupun nasional.

Kata Kunci: ***Indeks Pembangunan Literasi Masyarakat, Bootstrap, Resampling, Literasi Masyarakat, IPLM.***

# 

# BAB I

**PENDAHULUAN**

## Latar Belakang

Literasi masyarakat merupakan salah satu indikator utama dalam menilai kualitas sumber daya manusia dan tingkat pembangunan suatu negara. Kemampuan individu untuk membaca, menulis, memahami informasi, serta memanfaatkan pengetahuan dalam kehidupan sehari-hari menjadi pondasi penting bagi peningkatan produktivitas, pengambilan keputusan yang tepat, dan partisipasi aktif dalam kehidupan sosial dan ekonomi [1]. Di era globalisasi dan digitalisasi saat ini, literasi tidak hanya terbatas pada kemampuan membaca dan menulis konvensional, tetapi juga mencakup literasi digital, literasi informasi, dan literasi media. Peningkatan literasi secara menyeluruh berkontribusi pada pemberdayaan masyarakat, peningkatan kesejahteraan, serta pengurangan kesenjangan sosial dan ekonomi antarwilayah. Di Indonesia, diketahui bahwa telah terdapat berbagai program pemerintah telah diluncurkan untuk meningkatkan literasi masyarakat [2].

Meskipun demikian, data survei terbaru menunjukkan adanya ketimpangan signifikan dalam tingkat literasi, baik antarprovinsi, wilayah perkotaan dan pedesaan, maupun antar kelompok sosial dan ekonomi. Ketimpangan ini menunjukkan bahwa pengukuran indeks literasi yang akurat menjadi hal penting agar kebijakan dan program literasi dapat lebih tepat sasaran. Pengukuran indeks pembangunan literasi masyarakat menghadapi tantangan dalam hal kualitas data, ukuran sampel, dan distribusi data yang sering kali tidak mengikuti asumsi normal. Banyak survei literasi menggunakan sampel terbatas atau data yang *skewed*, sehingga hasil perhitungan awal dapat memiliki tingkat ketidakpastian yang cukup tinggi [3]. Oleh karena itu, diperlukan metode statistik yang mampu memperhitungkan ketidakpastian dan variasi data secara lebih akurat. Salah satu metode yang relevan adalah metode bootstrap, yakni teknik resampling non-parametrik.

Penerapan metode bootstrap dalam konteks indeks pembangunan literasi masyarakat memberikan sejumlah keuntungan. Pertama, metode ini mampu menghasilkan estimasi yang lebih stabil pada data yang jumlah sampelnya terbatas [4]. Kedua, metode bootstrap memungkinkan analisis risiko kesalahan dan ketidakpastian estimasi secara empiris. Ketiga, hasil analisis dengan bootstrap dapat dijadikan dasar pengambilan keputusan berbasis data yang lebih handal untuk program literasi di masyarakat. Semakin berkembangnya literasi digital dan tersedianya data besar (*big data*) dari survei daring, metode bootstrap menjadi semakin relevan untuk memberikan gambaran indeks literasi lebih representatif.

Berdasarkan latar belakang tersebut, penelitian ini berfokus pada penerapan metode bootstrap untuk menganalisis indeks pembangunan literasi masyarakat Indonesia tahun 2024. Oleh sebab itu, dengan menggunakan pendekatan ini, diharapkan diperoleh pemahaman yang lebih mendalam tentang kondisi literasi masyarakat, distribusi hasil survei, serta tingkat kepercayaan terhadap estimasi indeks literasi. Penelitian ini menjadi penting karena hasilnya dapat menjadi acuan bagi pembuat kebijakan, peneliti, dan lembaga pendidikan dalam merancang strategi pengembangan literasi yang efektif, efisien, dan berbasis bukti.

## Rumusan Masalah

Berdasarkan uraian latar belakang di atas, rumusan masalah dalam penelitian ini:

1. Bagaimana penerapan metode bootstrap dalam analisis statistik untuk memperkirakan indeks pembangunan literasi masyarakat Indonesia pada tahun 2024?
2. Bagaimana metode bootstrap dapat membantu mengatasi keterbatasan data atau distribusi yang tidak normal dalam pengukuran indeks literasi?
3. Bagaimana interpretasi hasil analisis bootstrap terhadap indeks pembangunan literasi dapat memberikan masukan bagi kebijakan literasi masyarakat?

## Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengevaluasi efektivitas metode bootstrap dalam menghasilkan estimasi Indeks Pembangunan Literasi Masyarakat (IPLM) dengan membandingkannya terhadap hasil estimasi yang diperoleh tanpa menggunakan dan menggunakan metode bootstrap.
2. Menyediakan gambaran distribusi indeks literasi masyarakat berdasarkan sampel survei tahun 2024, lengkap dengan interval kepercayaan.
3. Memberikan arahan mengenai potensi penggunaan metode bootstrap sebagai pendekatan analitis untuk mendukung proses perencanaan dan pengembangan program literasi masyarakat di berbagai wilayah.

## Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

### Manfaat Praktis:

1. Memberikan data dan estimasi indeks literasi masyarakat yang lebih akurat bagi pembuat kebijakan, lembaga pendidikan, dan organisasi non-pemerintah.
2. Membantu menentukan prioritas intervensi literasi, terutama untuk daerah atau kelompok sosial yang membutuhkan perhatian lebih.

### Manfaat Akademis:

1. Memberikan kontribusi pada pengembangan metodologi statistik, khususnya penerapan metode bootstrap dalam penelitian sosial dan survei literasi.
2. Menjadi referensi bagi peneliti lain dalam menggunakan teknik resampling non-parametrik untuk analisis data sosial.

# BAB II

**LANDASAN TEORI**

## Teknik Resampling

Teknik resampling merupakan sekelompok metode statistik yang melakukan pengambilan sampel berulang kali dari data yang tersedia untuk melakukan inferensi statistik. Resampling menjadi alternatif penting ketika asumsi-asumsi statistik parametrik tidak terpenuhi atau ketika distribusi sampling dari suatu statistik sulit ditentukan secara teoritis [5]. Metode ini tidak bergantung pada asumsi distribusi tertentu, sehingga disebut juga sebagai metode nonparametrik atau *distribution-free method* [6]. Prinsip dasar teknik resampling adalah melakukan pengambilan sampel berulang dari data yang tersedia untuk membentuk sejumlah sampel baru, sehingga karakteristik distribusi dari statistik yang diukur dapat diestimasi secara empiris [7].

Teknik resampling memiliki keunggulan utama karena tidak memerlukan asumsi distribusi tertentu pada data, sehingga sangat berguna ketika distribusi populasi tidak diketahui atau data tidak memenuhi asumsi parametrik konvensional seperti normalitas [7]. Dengan menggunakan teknik resampling, peneliti dapat melakukan inferensi statistik yang lebih robust dan reliable tanpa harus memenuhi asumsi-asumsi distribusi yang rumit, sehingga metode ini sangat cocok untuk data yang kompleks atau tidak memenuhi asumsi normalitas [6].

Resampling dapat dilakukan dengan dua pendekatan utama yaitu resampling dengan pengembalian (*bootstrap*) dan resampling tanpa pengembalian (*jackknife*), dimana metode bootstrap melakukan pengambilan sampel secara *resampling with replacement* untuk mendapatkan kesalahan standar pada pengujian hipotesis. Penelitian tentang estimasi *confidence interval* menggunakan metode bootstrap pada regresi komponen utama menunjukkan bahwa metode resampling memberikan fleksibilitas yang lebih besar dalam asumsi distribusi dibandingkan dengan metode klasik, meskipun *confidence interval* yang dihasilkan tidak berbeda signifikan dengan metode konvensional [7]. Ukuran resampling yang lebih besar akan menghasilkan distribusi yang lebih stabil dan estimasi yang lebih akurat, namun juga memerlukan waktu komputasi yang lebih lama. Penelitian tentang pengaplikasian metode bootstrap dalam membangun selang kepercayaan prediksi menerapkan teknik resampling bootstrap dengan pengulangan residual sebanyak B = 500, 1000, dan 5000 kali dan hasilnya menunjukkan bahwa semakin banyak pengulangan resampling maka semakin stabil distribusi bootstrap yang dihasilkan [8].

Teknik resampling sangat berguna dalam situasi di mana rumus analitik untuk standar error atau distribusi sampling tidak tersedia atau sangat kompleks untuk diturunkan [5]. Beberapa metode resampling yang umum digunakan dalam analisis statistik meliputi *bootstrap, jackknife*, *permutation test*, dan *cross-validation*, di mana setiap metode memiliki karakteristik dan aplikasi yang berbeda tergantung pada tujuan analisis dan jenis data yang digunakan [6]. Metode bootstrap adalah prosedur statistik nonparametrik yang digunakan untuk menguji signifikansi dalam berbagai analisis statistik, dengan melakukan pengambilan sampel secara *resampling with replacement* untuk mendapatkan kesalahan standar pada pengujian hipotesis [5]. Penelitian yang dilakukan oleh *Fallo et al* menunjukkan bahwa metode bootstrap efektif digunakan untuk memperoleh distribusi Indeks Moran dan nilai-p dalam analisis spasial kebutuhan pangan di Indonesia, dengan hasil yang menunjukkan adanya autokorelasi spasial yang signifikan [9]. Dalam praktiknya, pemilihan metode resampling yang tepat harus mempertimbangkan karakteristik data, ukuran sampel, dan tujuan analisis yang ingin dicapai [10].

Dalam konteks analisis data IPLM, teknik resampling dapat memberikan estimasi yang lebih akurat tentang variabilitas statistik deskriptif dan membantu dalam membuat inferensi yang lebih reliable, terutama ketika ukuran sampel terbatas atau distribusi data tidak diketahui dengan pasti [9]. Keunggulan utama teknik resampling adalah kemampuannya untuk memberikan solusi praktis terhadap masalah inferensi statistik tanpa memerlukan asumsi distribusi yang ketat. Metode ini sangat fleksibel dan dapat diterapkan pada berbagai jenis data dan statistik, mulai dari yang sederhana seperti *mean* dan *median* hingga yang kompleks seperti koefisien regresi dan rasio [6].

Penelitian komparatif menunjukkan bahwa metode bootstrap merupakan metode yang efisien dibandingkan metode jackknife, hal ini didukung dengan kecilnya tingkat MSE yang dihasilkan dalam estimasi kurtosis dan skewness [6]. Penggunaan teknik resampling dalam analisis IPLM memungkinkan peneliti untuk memperoleh estimasi parameter yang lebih akurat dan *confidence interval* yang lebih reliable, sehingga kesimpulan yang diambil memiliki validitas yang lebih tinggi [9]. Dengan demikian, teknik resampling memberikan kontribusi penting dalam meningkatkan kualitas analisis statistik, terutama untuk data survei dan indeks komposit seperti IPLM yang memiliki struktur data yang kompleks dan seringkali tidak memenuhi asumsi-asumsi statistik parametrik [5].

## Bootstrap

Bootstrap adalah salah satu metode resampling yang paling populer dan banyak digunakan dalam analisis statistik modern. Prinsip dasar bootstrap adalah melakukan pengambilan sampel secara acak dengan pengembalian (random sampling with replacement) dari data asli untuk membentuk distribusi sampling empiris [6]. Melalui proses resampling berulang, biasanya dengan jumlah iterasi antara 1000 hingga 10000 kali, bootstrap dapat menghasilkan estimasi yang akurat untuk berbagai parameter statistik.

Secara matematis, proses bootstrap dapat dijelaskan sebagai berikut. Misalkan terdapat sampel asli X = {x₁, x₂, ..., xₙ} dengan ukuran n, maka dilakukan pengambilan sampel bootstrap X\* = {x₁, x₂, ..., xₙ} dengan pengembalian sebanyak B kali iterasi. Untuk setiap sampel bootstrap, dihitung statistik yang diminati θ̂, sehingga diperoleh distribusi bootstrap   
. Distribusi empiris yang terbentuk dari sampel-sampel bootstrap ini kemudian digunakan untuk melakukan inferensi statistik, termasuk estimasi standar error, bias, dan interval kepercayaan [11].

Estimator bootstrap untuk *mean* dapat dinyatakan sebagai:

dimana adalah estimasi *mean* untuk sampel bootstrap ke-b, dan adalah observasi ke-i dalam sampel bootstrap ke-b [6].

Standar error bootstrap dapat dihitung menggunakan formula:

dimana adalah rata-rata dari semua estimasi bootstrap [11].

Terdapat beberapa metode untuk menghitung confidence interval menggunakan bootstrap, antara lain metode percentile, metode bias-corrected and accelerated (BCa), dan metode bootstrap-t. Metode percentile adalah yang paling sederhana dan paling sering digunakan, dimana interval kepercayaan (1-α)100% diperoleh dengan mengambil percentile dan dari distribusi bootstrap. Interval kepercayaan percentile dapat dinyatakan sebagai:

dimana adalah percentile dari distribusi bootstrap dan adalah percentile [6].

Hasil penelitian *Sihombing dan Simamora* yang menggunakan data kekuatan gempa bumi di Indonesia tahun 2020 dengan magnitude di atas 5, menunjukkan bahwa metode bootstrap menghasilkan nilai *Mean Square Error* (MSE) yang lebih kecil dibandingkan metode jackknife [6]. Penelitian tersebut melakukan resampling sebanyak 50, 100, 200, 500, dan 1000 kali, dan menyimpulkan bahwa metode bootstrap merupakan metode yang lebih efisien dalam estimasi kurtosis dan skewness dibandingkan metode jackknife, hal ini didukung dengan kecilnya tingkat MSE yang dihasilkan.

Bootstrap memiliki berbagai keunggulan dibandingkan metode inferensi klasik, terutama dalam hal tidak memerlukan asumsi distribusi normal dan dapat diterapkan pada berbagai jenis statistik, termasuk yang tidak memiliki formula analitis untuk standar error. Metode ini sangat robust terhadap outlier dan dapat memberikan estimasi yang akurat bahkan untuk ukuran sampel yang relatif kecil [12]. Selain itu, bootstrap juga dapat digunakan untuk menilai stabilitas dan keandalan estimasi parameter dalam berbagai konteks penelitian.

## Penilaian Berdasarkan Statistika Deskriptif: *Mean*, *Median*, Standar Deviasi

Statistika deskriptif merupakan metode analisis data yang penting untuk menggambarkan dan menganalisis karakteristik utama data secara sistematis, dengan tahapan yang mencakup persiapan data, pengolahan, analisis, visualisasi, dan interpretasi hasil [13]. *Sunita & Syakur* menekankan bahwa penerapan statistik deskriptif seringkali menghadapi tantangan yang berpotensi mengurangi validitas hasil penelitian, sehingga diperlukan panduan praktis penerapan statistik deskriptif dalam penelitian kuantitatif dan kualitatif yang mencakup langkah-langkah utama dalam persiapan data, pengolahan, analisis, visualisasi, dan interpretasi hasil [13].

*Sa'adah* dalam penelitiannya mengenai analisis distribusi dan tren nilai mahasiswa mata kuliah Kalkulus periode 2019-2023 menjelaskan bahwa statistik deskriptif digunakan untuk menganalisis data dengan cara mendeskripsikan karakteristik data melalui perhitungan ukuran pemusatan seperti *mean* dan *median*, serta ukuran penyebaran seperti standar deviasi, dimana hasil analisis menunjukkan bahwa nilai rata-rata (*mean*) mahasiswa berada pada kategori menengah dengan standar deviasi yang bervariasi setiap tahun, misalnya tahun 2021 memiliki rata-rata terendah (50,65) dengan standar deviasi tertinggi (21,56) yang menunjukkan sebaran nilai cukup lebar [14] .

### *Mean*

*Mean* atau rata-rata adalah ukuran pemusatan data yang paling umum digunakan dan merupakan nilai yang mewakili himpunan atau sekelompok data. *Mean* dihitung dengan menjumlahkan seluruh nilai data kemudian membaginya dengan banyaknya data [15]. Secara matematis, *mean* untuk data sampel dapat dinyatakan sebagai:

dimana x̄ adalah *mean* sampel, xᵢ adalah nilai data ke-i, dan n adalah jumlah data [11].

*Mean* sangat sensitif terhadap nilai ekstrim atau outlier dalam dataset. Ketika terdapat nilai yang sangat berbeda dari nilai-nilai lainnya, *mean* dapat memberikan representasi yang kurang akurat terhadap pusat data. Namun demikian, *mean* tetap menjadi ukuran yang paling banyak digunakan karena kemudahan perhitungan dan interpretasinya, serta banyak digunakan dalam analisis statistik lanjutan [15].

### *Median*

*Median* adalah nilai tengah dari data yang telah diurutkan dari yang terkecil hingga terbesar. *Median* membagi data menjadi dua bagian yang sama banyak, sehingga 50% data berada di bawah *median* dan 50% data berada di atas *median* [11]. Keuntungan utama *median* adalah tidak terpengaruh oleh nilai ekstrim, sehingga memberikan ukuran pemusatan yang lebih robust dibandingkan *mean* untuk data yang memiliki outlier.

Untuk data dengan jumlah ganjil, *median* adalah nilai yang berada di posisi tengah setelah data diurutkan. Untuk data dengan jumlah genap, *median* adalah rata-rata dari dua nilai tengah [15]. Dalam analisis data yang memiliki distribusi yang sangat miring (*skewed*), *median* sering kali lebih representatif dibandingkan *mean* dalam menggambarkan kecenderungan sentral data.

### Standar Deviasi

Standar deviasi merupakan ukuran penyebaran data yang menunjukkan seberapa jauh data menyebar dari nilai rata-ratanya. Standar deviasi diperoleh dengan mengakar kuadratkan nilai varians dan memberikan informasi tentang variabilitas atau keragaman data [15]. Semakin besar nilai standar deviasi, maka data semakin menyebar dari nilai *mean*, sebaliknya jika nilai standar deviasi semakin kecil, maka data semakin terkonsentrasi di sekitar nilai *mean*.

Standar deviasi untuk data sampel dapat dihitung dengan rumus:

(5)

dimana s adalah standar deviasi sampel, xᵢ adalah nilai data ke-i, x̄ adalah *mean* sampel, dan n adalah jumlah data [11].

Standar deviasi sangat penting dalam analisis bootstrap karena digunakan untuk menilai konsistensi dan stabilitas estimasi parameter. Perbandingan standar deviasi sebelum dan sesudah proses bootstrap dapat memberikan informasi tentang efektivitas metode resampling dalam menghasilkan estimasi yang lebih reliable [6]. Dalam konteks penelitian, standar deviasi juga digunakan untuk menilai variabilitas data dan membandingkan tingkat keragaman antar kelompok atau sampel yang berbeda [12].

## Indeks Pembangunan Literasi Masyarakat

Indeks Pembangunan Literasi Masyarakat (IPLM) adalah pengukuran terhadap usaha yang dilaksanakan oleh pemerintah daerah (tingkat provinsi dan kabupaten/kota) dalam membina dan mengembangkan perpustakaan sebagai wahana belajar sepanjang hayat untuk mencapai budaya literasi masyarakat. IPLM merupakan data tingkat pembangunan literasi masyarakat yang diperoleh dari unsur-unsur pembangun literasi masyarakat (UPLM) yang bersumber dari data sekunder dan aspek masyarakat (AM) dalam upaya membina dan mengembangkan perpustakaan sebagai wahana belajar sepanjang hayat guna meningkatkan literasi masyarakat [16]. Indeks Pembangunan Literasi Masyarakat (IPLM) merupakan indikator penting dalam mengukur tingkat literasi masyarakat di Indonesia yang dikembangkan dan dipublikasikan oleh Perpustakaan Nasional Republik Indonesia sebagai alat evaluasi kemajuan literasi di seluruh wilayah Indonesia [17].

### Perkembangan IPLM di Indonesia

Berdasarkan data dari Perpustakaan Nasional Republik Indonesia, IPLM nasional tahun 2024 mencatat pencapaian luar biasa dengan skor 73,52 melampaui target 71,4 dan hasil tahun 2023 yang berada di angka 69,42, capaian ini menunjukkan peningkatan yang signifikan dalam literasi masyarakat Indonesia dan menunjukkan keberhasilan berbagai program pemberdayaan literasi yang telah dilaksanakan [17]. IPLM mencakup tujuh unsur penting yaitu pemerataan layanan perpustakaan, ketercukupan koleksi, tingkat kunjungan masyarakat, dan aspek lainnya, dimana kajian IPLM melibatkan 514 kabupaten/kota dan lebih dari 174 ribu responden usia 10-69 tahun sehingga menghasilkan dataset yang besar dan kompleks yang memerlukan analisis statistik yang robust untuk menghasilkan estimasi yang akurat [17]. Berdasarkan data Badan Pusat Statistik, tingkat capaian pembangunan literasi masyarakat Indonesia mengalami peningkatan dari tahun ke tahun [16]. Pada tahun 2022, nilai IPLM nasional berada pada angka 64,40, kemudian pada tahun 2023 meningkat menjadi 64,68, menunjukkan peningkatan sebesar 1,03 poin [18].

Pada tahun 2024, IPLM nasional mencapai 73,52 poin, menandakan peningkatan yang signifikan dalam upaya pembangunan literasi masyarakat di Indonesia. Kategori penilaian IPLM terdiri dari lima tingkatan, yaitu: sangat rendah (0-29,9), rendah (30-49,9), sedang (50-79,9), tinggi (80-89,9), dan sangat tinggi (90-100) [17]. Meskipun telah terjadi peningkatan, capaian IPLM Indonesia masih berada pada kategori sedang, yang menunjukkan masih diperlukan upaya yang lebih intensif untuk meningkatkan literasi masyarakat di seluruh wilayah Indonesia.

### IPLM di Pulau Sumatera

Wilayah Sumatera memiliki karakteristik pembangunan literasi yang bervariasi antar provinsi. Berdasarkan data IPLM tahun 2024, beberapa provinsi di Sumatera menunjukkan pencapaian yang cukup baik, dengan Sumatera Barat berada di peringkat ke-4 nasional dengan nilai 82,47 poin, masuk dalam kategori tinggi [16]. Provinsi lainnya seperti Kepulauan Riau dan Jambi juga menunjukkan nilai IPLM yang berada pada kategori sedang hingga tinggi.

Perbedaan capaian IPLM antar wilayah di Sumatera mencerminkan tantangan dalam pemerataan pembangunan literasi. Data IPLM dari kabupaten dan kota yang tersebar di seluruh provinsi di Sumatera pada tahun 2024 memberikan gambaran komprehensif tentang kondisi literasi masyarakat di wilayah ini [16]. Analisis terhadap data ini menjadi penting untuk merumuskan kebijakan yang tepat dalam meningkatkan pembangunan literasi di tingkat daerah, terutama dalam konteks disparitas regional yang masih cukup tinggi di wilayah Sumatera.

# BAB III

**METODE PENELITIAN**

## Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan deskriptif dan inferensial. Pendekatan ini digunakan untuk menggambarkan karakteristik data Indeks Pembangunan Literasi Masyarakat (IPLM) serta menerapkan metode bootstrap sebagai teknik resampling untuk memperoleh estimasi statistik yang lebih akurat dan robust. Penelitian kuantitatif dipilih karena fokus utama analisis berada pada pengukuran, perhitungan numerik, dan penarikan kesimpulan statistik berdasarkan data survei yang tersedia.

## Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) dan Perpustakaan Nasional Republik Indonesia. Data tersebut berisi nilai Indeks Pembangunan Literasi Masyarakat (IPLM) tahun 2024 untuk seluruh kabupaten/kota di Pulau Sumatera. Data sekunder dipilih karena telah melalui proses pengumpulan, verifikasi, dan publikasi resmi, sehingga valid untuk digunakan dalam analisis statistik.

## Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan dengan dokumentasi, yaitu mengunduh dataset resmi dari situs BPS dan laporan IPLM dari Perpustakaan Nasional RI. Data kemudian diseleksi, diverifikasi, serta dibersihkan (data cleaning) untuk memastikan tidak terdapat nilai hilang (missing values) atau ketidaksesuaian format yang dapat mengganggu proses analisis.

## Variabel yang Diamati

Variabel yang diamati dalam penelitian ini merujuk pada komponen atau karakteristik tertentu yang diukur untuk memperoleh informasi terkait fenomena yang diteliti. Pada konteks penelitian ini, variabel yang diamati mencakup data yang merepresentasikan kondisi pembangunan literasi masyarakat di wilayah Sumatera tahun 2024. Pengamatan terhadap variabel ini dilakukan untuk memahami variasi nilai antar wilayah, mengidentifikasi pola distribusi data, serta mendukung proses analisis statistik yang akan dilakukan.

**Tabel 3.1** Operasionalisasi Variabel

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variabel** | **Definisi** | **Simbol** | **Satuan** | **Skala** |
| (1) | (2) | (3) | (4) | (5) |
| Indeks Pembangunan Literasi Masyarakat (IPLM) | Ukuran capaian pembangunan literasi masyarakat yang menggambarkan tingkat kemampuan masyarakat dalam mengakses, memahami, dan memanfaatkan informasi | IPLM | Skor Indeks | Rasio |
| Kabupaten/Kota | Unit wilayah administratif tempat data diperoleh | KK | Nama Daerah | Nominal |

Berdasarkan Tabel 3.1 Operasionalisasi Variabel, penelitian ini menggunakan dua jenis variabel, yaitu variabel utama dan variabel pendukung. Variabel utama dalam penelitian ini adalah Indeks Pembangunan Literasi Masyarakat (IPLM), yang merepresentasikan tingkat capaian literasi masyarakat melalui kemampuan dalam mengakses, memahami, serta memanfaatkan informasi. Variabel ini bersifat kuantitatif dengan satuan skor indeks dan menggunakan skala rasio sehingga memungkinkan dilakukan analisis statistik inferensial. Selain itu, penelitian ini juga menggunakan Kabupaten/Kota sebagai variabel pendukung, yang berfungsi untuk mengidentifikasi wilayah administratif tempat data diperoleh. Variabel ini bersifat kategorik dengan satuan berupa nama daerah serta menggunakan skala nominal. Penggunaan variabel ini memungkinkan peneliti untuk meninjau distribusi serta perbandingan capaian IPLM antar wilayah.

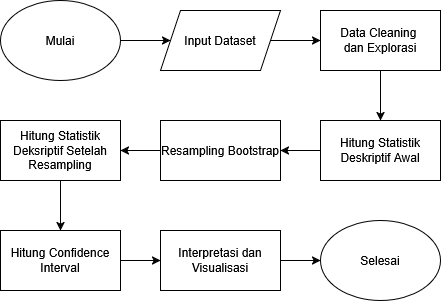
**Tabel 3.2** Dataset Penelitian

|  |  |  |
| --- | --- | --- |
| **X1** | **X1** | **X1** |
| (1) | (2) | (3) |
| No | Kabupaten/Kota | IPLM\_2024 |
| 1 | Simeulue | 45.43 |
| 2 | Aceh Singkil | 58.41 |
| **Tabel 3.2** Dataset Penelitian (Lanjutan) | | |
| **X1** | **X1** | **X1** |
| (1) | (2) | (3) |
| No | Kabupaten/Kota | IPLM\_2024 |
| 3 | Aceh Selatan | 47.87 |
| 4 | Aceh Tenggara | 47.5 |
| 5 | Aceh Timur | 44.47 |
| 6 | Aceh Tengah | 81.65 |
| 7 | Aceh Barat | 55.37 |
| … | … | … |
| 129 | Pesawaran | 53.83 |
| 130 | Pringsewu | 60.25 |
| 131 | Mesuji | 60.37 |
| 132 | Tulang Bawang Barat | 60.68 |
| 133 | Pesisir Barat | 50.81 |
| 134 | Kota Bandar Lampung | 76.74 |
| 135 | Kota Metro | 94.41 |

Berdasarkan Tabel 3.2 Dataset Penelitian, Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang berisi nilai Indeks Pembangunan Literasi Masyarakat (IPLM) tahun 2024 pada tingkat kabupaten dan kota di wilayah Sumatera. Data ini mencakup informasi mengenai wilayah administratif sebagai identitas lokasi serta skor IPLM sebagai indikator capaian pembangunan literasi masyarakat. Penyajian dataset ini bertujuan untuk memberikan gambaran awal mengenai distribusi nilai IPLM di berbagai daerah, sehingga dapat digunakan sebagai dasar dalam proses analisis statistik deskriptif maupun penerapan metode bootstrap pada tahap penelitian selanjutnya.

## Diagram Alir

Diagram alir digunakan dalam penelitian ini untuk memberikan gambaran visual mengenai tahapan proses analisis data yang dilakukan. Penyajian diagram alir membantu memperjelas urutan langkah kerja mulai dari pengumpulan data, pengolahan, analisis, dan pengambilan keputusan ataupun kesimpulan.



Gambar 3.1 Diagram Alir

Berdasarkan Gambar 3.1, diagram alir tersebut menggambarkan rangkaian proses analisis data menggunakan pendekatan bootstrap secara sistematis. Proses dimulai pada tahap input dataset, yaitu memasukkan data IPLM yang menjadi objek penelitian. Selanjutnya dilakukan Data cleaning dan eksplorasi untuk memastikan bahwa data yang digunakan berada dalam kondisi layak analisis, bebas dari kesalahan pencatatan, serta memberikan gambaran awal mengenai karakteristik data. Setelah tahap pra-pemrosesan tersebut, dilakukan perhitungan statistik deskriptif swal guna mengetahui nilai *mean*, *median*, dan standar deviasi sebelum resampling.

Tahap berikutnya adalah resampling bootstrap, yaitu proses pengambilan sampel ulang dengan pengembalian untuk menghasilkan sejumlah besar sampel tiruan. Sampel-sampel ini digunakan untuk menghitung kembali statistik deskriptif setelah resampling sehingga diperoleh estimasi yang lebih stabil dan tidak terlalu bergantung pada satu sampel awal. Setelah itu, dilakukan perhitungan confidence interval untuk menentukan batas bawah dan batas atas estimasi dengan tingkat kepercayaan tertentu. Pada tahap akhir, hasil analisis kemudian diinterpretasikan dan divisualisasikan agar temuan dapat disajikan secara jelas dan informatif. Seluruh rangkaian proses diakhiri pada tahap selesai. Dengan mengikuti alur ini, analisis bootstrap dapat dilakukan secara terstruktur dan menghasilkan estimasi statistik yang lebih reliabel.

## Tools dan Software

Analisis data dalam penelitian ini dilakukan berbantuan perangkat lunak RStudio dengan basis bahasa pemrograman R. Software ini dipilih karena memiliki tools yang dapat menangani metode statistik baik sederhana maupun kompleks. ara spesifik, proses analisis didukung oleh sejumlah paket pustaka (*libraries*) dengan fungsi tertentu. Adapun daftar *library* yang digunakan dalam penelitian ini adalah sebagai berikut:

1. readxl: Digunakan pada tahap awal untuk mengimpor dataset sumber (.xlsx) ke dalam lingkungan kerja R.
2. dplyr: Digunakan untuk manajemen data (*data wrangling*), termasuk seleksi variabel dan transformasi data sebelum analisis.
3. boot: Merupakan instrumen utama untuk menjalankan prosedur *resampling* (bootstrap) guna menghasilkan estimasi parameter *mean*, *median*, dan standar deviasi yang kokoh.
4. ggplot2: Digunakan untuk menyusun visualisasi data yang informatif, termasuk pembuatan grafik *Error Bar* untuk interval kepercayaan.

# BAB IV

**HASIL DAN PEMBAHASAN**

## Statistik Deskriptif Sebelum Bootstrap

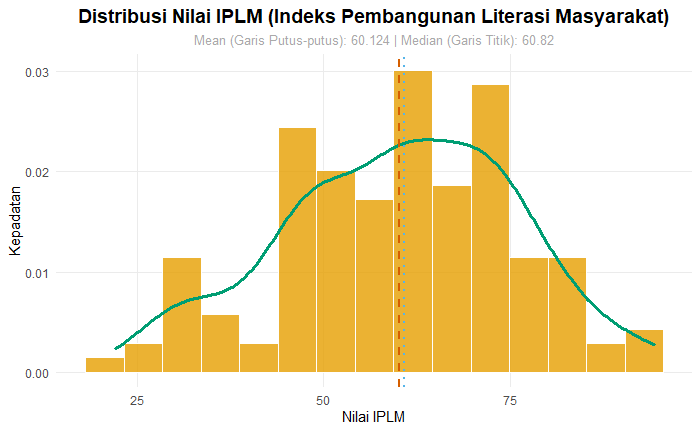
Analisis statistik deskriptif digunakan sebagai tahap awal untuk memahami karakteristik umum dari data Indeks Pembangunan Literasi Masyarakat (IPLM) 2024. Tahap ini bertujuan untuk menggambarkan pola sebaran nilai IPLM sebelum dilakukan proses resampling menggunakan metode bootstrap. Statistik deskriptif memberikan informasi dasar seperti ukuran pemusatan (*mean*, *median*) dan ukuran penyebaran (standar deviasi), sehingga peneliti dapat mengidentifikasi kecenderungan umum dan tingkat variasi antarwilayah. Pemahaman awal ini penting karena kualitas distribusi data akan berpengaruh pada hasil estimasi bootstrap, termasuk stabilitas nilai *mean*, *median*, dan interval kepercayaan. Dengan demikian, tahap statistik deskriptif awal berfungsi sebagai acuan pembanding untuk menilai perubahan nilai setelah proses bootstrap dilakukan.

**Tabel 4.1.** Statistik Deskriptif Sebelum Bootstrap

|  |  |
| --- | --- |
| **Statistik** | **Nilai** |
| (1) | (2) |
| *Mean* | 60.12356 |
| *Median* | 60.82 |
| Standar Deviasi | 15.68139 |

Berdasarkan hasil perhitungan, nilai *mean* IPLM sebesar 60.12356, yang menunjukkan bahwa rata-rata capaian literasi tingkat wilayah berada pada kisaran skor 60 dari skala 0–100. Nilai ini menggambarkan kondisi pembangunan literasi nasional yang berada pada kategori sedang, namun masih memiliki ruang peningkatan pada beberapa wilayah. Nilai *median* sebesar 60.82, yang sedikit lebih tinggi daripada *mean*, mengindikasikan bahwa distribusi data relatif seimbang namun cenderung memiliki beberapa nilai rendah yang menarik rata-rata ke bawah. Sementara itu, standar deviasi yang mencapai 15.68139 menunjukkan adanya variasi yang cukup besar antarwilayah dalam pencapaian IPLM. Variabilitas ini menandakan bahwa terdapat ketimpangan kemampuan literasi antara wilayah dengan performa tinggi dan wilayah yang masih tertinggal.

Berdasarkan hasil perhitungan, nilai *mean* IPLM sebesar 60.12356, yang menunjukkan bahwa rata-rata capaian literasi tingkat wilayah berada pada kisaran skor 60 dari skala 0–100. Nilai ini menggambarkan kondisi pembangunan literasi nasional yang berada pada kategori sedang, namun masih memiliki ruang peningkatan pada beberapa wilayah. Nilai *median* sebesar 60.82, yang sedikit lebih tinggi daripada *mean*, mengindikasikan bahwa distribusi data relatif seimbang namun cenderung memiliki beberapa nilai rendah yang menarik rata-rata ke bawah. Sementara itu, standar deviasi yang mencapai 15.68139 menunjukkan adanya variasi yang cukup besar antarwilayah dalam pencapaian IPLM. Variabilitas ini menandakan bahwa terdapat ketimpangan kemampuan literasi antara wilayah dengan performa tinggi dan wilayah yang masih tertinggal.



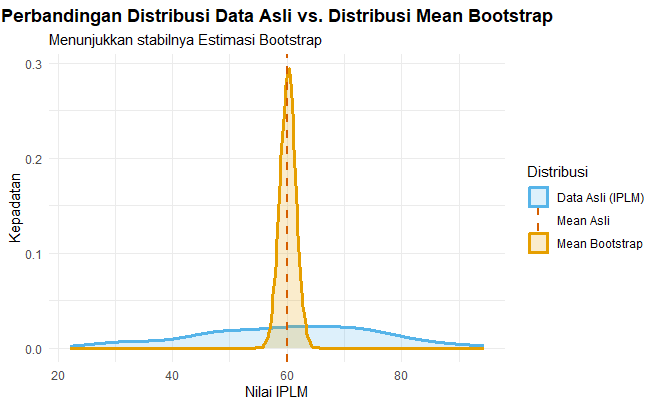
Gambar 4.1 Sebaran Nilai IPLM 2024 Berdasarkan Data Asli

Berdasarkan grafik distribusi pada Gambar 4.1, terlihat bahwa nilai IPLM tidak berdistribusi normal sempurna dan memiliki penyebaran yang cukup lebar antarwilayah. Histogram menunjukkan bahwa sebagian besar nilai IPLM berada pada kisaran 55–75, sementara kurva kepadatan memperlihatkan sedikit ketidaksimetrisan. Garis vertikal putus-putus menandai nilai *mean* sebesar 60.12, dan garis titik-titik menandai *median* sebesar 60.82. Perbedaan kecil antara *mean* dan *median* mengindikasikan bahwa distribusi data relatif seimbang tetapi masih dipengaruhi beberapa nilai ekstrem yang menarik rata-rata sedikit ke bawah. Selain itu, standar deviasi sebesar 15.68 menegaskan adanya variasi yang cukup tinggi antarwilayah, mencerminkan ketimpangan capaian literasi di berbagai daerah. Informasi ini penting dasar pembanding untuk proses bootstrap, karena metode tersebut akan mengevaluasi kembali stabilitas estimasi *mean*, *median*, dan varians berdasarkan distribusi sampel ulang.

## Proses Resampling Bootstrap

Proses resampling bootstrap dilakukan untuk memperoleh estimasi distribusi sampling dari statistik deskriptif Indeks Pembangunan Literasi Masyarakat (IPLM) 2024 secara lebih akurat dan robust. Teknik ini dipilih karena distribusi empiris IPLM antar-wilayah menunjukkan ketidaksimetrisan dan variasi antar daerah yang cukup besar, sehingga penggunaan pendekatan analitis berbasis asumsi distribusi normal dianggap kurang tepat. Melalui bootstrap, estimasi parameter seperti *mean*, *median*, dan varians dihitung langsung dari pola variasi yang muncul dalam data tanpa bergantung pada bentuk distribusi tertentu.

Pada tahap awal, dataset IPLM 2024 yang berisi nilai indeks seluruh provinsi/kabupaten/kota digunakan sebagai dasar resampling. Setiap proses bootstrap dilakukan dengan mekanisme sampling *with replacement*, sehingga setiap observasi memiliki peluang terpilih lebih dari satu kali. Ukuran sampel pada setiap replikasi ditetapkan sama seperti dataset asli, dan dalam penelitian ini dilakukan 5.000 replikasi untuk menghasilkan estimasi distribusi statistik yang stabil. Hasil dari ribuan replikasi ini kemudian digunakan untuk membangun distribusi empiris berbagai statistik deskriptif. Distribusi inilah yang divisualisasikan pada grafik berikut, sehingga memberikan gambaran yang lebih jelas mengenai konsistensi estimasi bootstrap dibandingkan distribusi data aslinya.



Gambar 4.2 Visualisasi Distribusi Nilai IPLM Asli dan Distribusi *Mean* Bootstrap

Visualisasi pada Gambar 4.2 menunjukkan perbandingan antara distribusi data asli IPLM dengan distribusi *mean* hasil bootstrap. Kurva biru menggambarkan pola sebaran nilai IPLM yang relatif lebar, menandakan adanya variasi antar wilayah yang cukup besar. Sementara itu, distribusi *mean* bootstrap (kurva kuning) terlihat jauh lebih sempit dan terpusat kuat di sekitar titik *mean* asli. Pola ini menggambarkan bahwa estimasi rata-rata dari proses resampling sangat stabil meskipun dilakukan ribuan kali, nilai *mean* cenderung berkumpul pada rentang yang sama dan tidak banyak menyimpang dari *mean* data asli. Stabilitas distribusi *mean* bootstrap tersebut menunjukkan bahwa proses resampling berhasil mengurangi pengaruh fluktuasi data antar wilayah tanpa mengubah informasi utama yang terkandung dalam data asli. Dengan kata lain, bootstrap memperlihatkan bahwa estimasi parameter statistik (khususnya *mean*) cukup konsisten dan tidak sensitif terhadap variasi sampel. Hal ini mendukung penggunaan interval kepercayaan berbasis bootstrap, karena distribusi hasil resampling benar-benar mencerminkan ketidakpastian estimasi *mean* IPLM 2024 secara empiris. Grafik ini juga memperlihatkan bagaimana bootstrap membantu menegaskan “inti” dari distribusi parameter meskipun data asli memiliki variasi yang lebih luas.

## Statistik Deskriptif Setelah Bootstrap

Analisis statistik deskriptif setelah *bootstrap* dilakukan sebagai tahap evaluasi akhir untuk mendapatkan estimasi parameter populasi yang lebih akurat dan stabil. Berbeda dengan tahap awal, tujuan utama dari statistik deskriptif pada tahap ini adalah untuk menggambarkan karakteristik distribusi data hasil proses *resampling* berulang. Informasi ini memberikan gambaran mengenai kekokohan (*robustness*) dari nilai rata-rata (*mean*), nilai tengah (*median*), dan tingkat variasi (*standar deviasi*) yang telah dikoreksi dari potensi bias pada sampel tunggal. Dengan demikian, hasil ini merepresentasikan estimasi yang lebih mendekati parameter populasi yang sebenarnya dan menjadi landasan yang kuat untuk penarikan kesimpulan statistik.

**Tabel 4.2** Statistik Deskriptif Setelah Bootstrap

|  |  |
| --- | --- |
| **Statistik** | **Nilai** |
| (1) | (2) |
| *Mean* | 60.12722 |
| *Median* | 61.01635 |
| Standar Deviasi | 15.60152. |

Penilaian statistik deskriptif setelah dilakukan *bootstrap* menunjukkan estimasi nilai pusat yang stabil. Nilai *mean* diperoleh sebesar 60.12722, yang mengindikasikan bahwa rata-rata capaian (dalam asumsi skala 0–100) tetap berada pada kategori sedang. Nilai ini menunjukkan konsistensi rata-rata populasi yang terjaga setelah proses *resampling*. Nilai *median* tercatat sebesar 61.01635. Nilai *median* yang sedikit lebih tinggi dibandingkan nilai *mean* ini mengisyaratkan bahwa distribusi data hasil *bootstrap* cenderung menceng ke kiri (*negatively skewed*), di mana mayoritas data berkumpul di area nilai yang lebih tinggi, namun terdapat sejumlah nilai rendah yang menarik rata-rata ke bawah. Di sisi lain, Standar Deviasi sebesar 15.60152 memperlihatkan bahwa variabilitas data masih cukup lebar. Hal ini menegaskan bahwa meskipun nilai pusat telah terestimasi dengan stabil, keragaman atau ketimpangan antar unit pengamatan dalam populasi tetap signifikan.

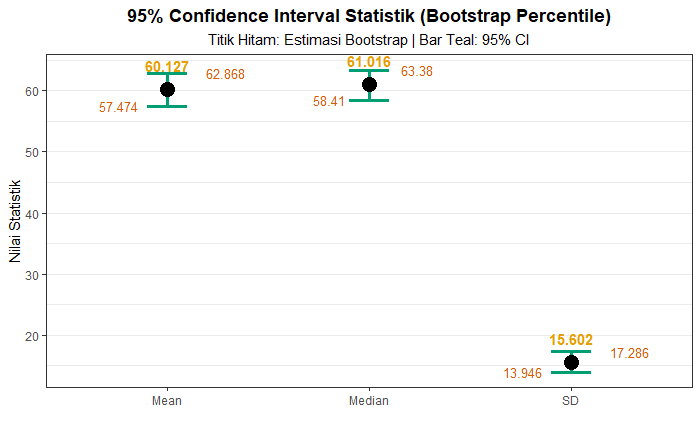
## Interval Kepercayaan Bootstrap

Analisis Interval Kepercayaan (*Confidence Interval*) 95% dengan metode persentil (*Percentile Method*) dilakukan untuk menilai tingkat presisi estimasi parameter statistik yang diperoleh melalui teknik *bootstrap*. Metode ini membentuk interval dengan mengambil nilai kuantil 2.5% sebagai batas bawah dan 97.5% sebagai batas atas dari distribusi *bootstrap*, sehingga memberikan rentang yang merepresentasikan posisi parameter populasi yang diperkirakan. Karena pendekatan ini tidak memerlukan asumsi distribusi tertentu, metode ini sangat efektif untuk data yang tidak berdistribusi normal atau memiliki ukuran sampel kecil. Dengan menggunakan selang kepercayaan 95%, peneliti dapat meyakini bahwa jika proses resampling dilakukan berulang kali, sekitar 95% interval yang dihasilkan akan memuat nilai parameter yang sebenarnya. Interval yang lebih sempit menunjukkan estimasi yang lebih presisi, sedangkan interval yang lebih lebar menandakan variabilitas yang lebih tinggi, sehingga informasi ini penting untuk menilai keandalan estimasi *mean*, *median*, dan standar deviasi yang telah dihitung sebelumnya.

**Tabel 4.3.** Confidence Interval 95% Percentile Method

|  |  |  |
| --- | --- | --- |
| **Statistik** | **2.5% Quantile** | **97.5 % Quantile** |
| (1) | (2) | (3) |
| *Mean* | 57.47392 | 62.86831 |
| *Median* | 58.41 | 63.38 |
| Standar Deviasi | 13.94621 | 17.28635 |

Berdasarkan parameter *mean*, interval kepercayaan terbentang antara 57.47 hingga 62.87. Hal ini mengindikasikan bahwa rata-rata capaian populasi yang sebenarnya sangat mungkin berada dalam kisaran tersebut, menegaskan stabilitas nilai rata-rata di sekitar angka 60. Selanjutnya, *Median* memiliki interval antara 58.41 hingga 63.38. Rentang yang beririsan kuat dengan interval *mean* ini menunjukkan bahwa distribusi data cenderung memusat dan konsisten. Sementara itu, untuk Standar Deviasi, interval estimasi berada pada kisaran 13.95 hingga 17.29. Batas-batas ini memberikan gambaran mengenai rentang variabilitas atau ketimpangan yang mungkin terjadi dalam populasi, di mana nilai variasi tidak akan jauh dari angka 15.6, memberikan kepastian mengenai tingkat heterogenitas data.

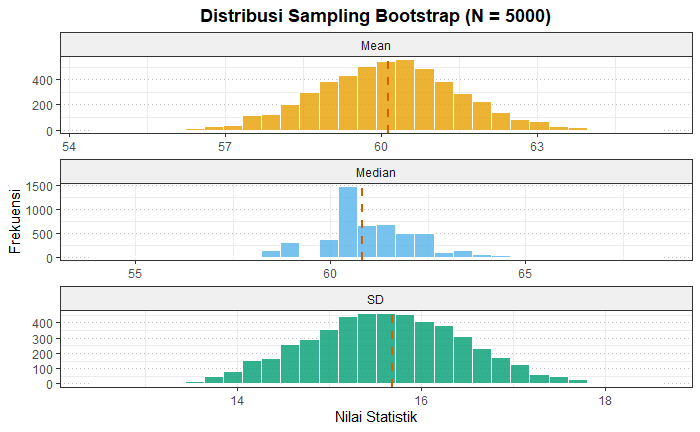


Gambar 4.3 Error Bar Confidence Interval

Gambar 4.3 menunjukkan dari rentang estimasi tersebut disajikan menggunakan grafik *Error Bar*. Pada grafik tersebut, titik hitam merepresentasikan estimasi titik (*point estimate*) untuk *mean*, *Median*, dan *Standar Deviasi*, sedangkan garis merah vertikal menggambarkan lebar interval kepercayaan (CI 95%). Secara visual, terlihat bahwa posisi *mean* dan *Median* berada pada ketinggian yang hampir setara dengan rentang *error bar* yang relatif pendek. Hal ini menandakan bahwa estimasi nilai pusat memiliki tingkat presisi yang tinggi. Sebaliknya, posisi *Error Bar* untuk *Standar Deviasi* berada jauh lebih rendah, yang memvisualisasikan besaran sebaran data secara terpisah. Grafik ini memudahkan pembacaan untuk melihat bahwa meskipun terdapat variasi dalam data yang ditunjukkan dalam *Standar Deviasi*, estimasi nilai rata-rata populasi tetap kokoh dan stabil dalam rentang yang sempit.

## Visualisasi Hasil Bootstrap

Visualisasi hasil bootstrap digunakan untuk memberikan gambaran yang lebih jelas mengenai distribusi statistik yang dihasilkan dari proses resampling sebanyak 5.000 kali. Pada penelitian ini, tiga statistik utama dianalisis melalui bootstrap, yaitu *mean*, *median*, dan standar deviasi Indeks Pembangunan Literasi Masyarakat (IPLM) tahun 2024. Ketiga distribusi bootstrap ini ditampilkan dalam bentuk histogram yang dilengkapi dengan garis vertikal menunjukkan nilai asli (sebelum bootstrap), sehingga memudahkan pembaca dalam membandingkan estimasi awal dengan hasil resampling.



Gambar 4.4 Distribusi Bootstrap untuk *Mean*, *Median*, dan Standard Deviation

Gambar 4.4 menunjukkan tiga histogram: *Bootstrap Distribution of Mean*, *Bootstrap Distribution of Median*, dan *Bootstrap Distribution of SD*. Pada histogram distribusi *mean*, tampak bahwa hasil bootstrap membentuk distribusi yang relatif simetris dan mendekati normal, dengan puncak berada di sekitar nilai 60. Hal ini menandakan bahwa estimasi rerata IPLM cukup stabil, dan resampling tidak menghasilkan variasi ekstrem. Garis merah di tengah menunjukkan nilai *mean* awal yang berada dekat pusat distribusi, mengindikasikan bahwa proses bootstrap tidak mengubah nilai estimasi secara signifikan.

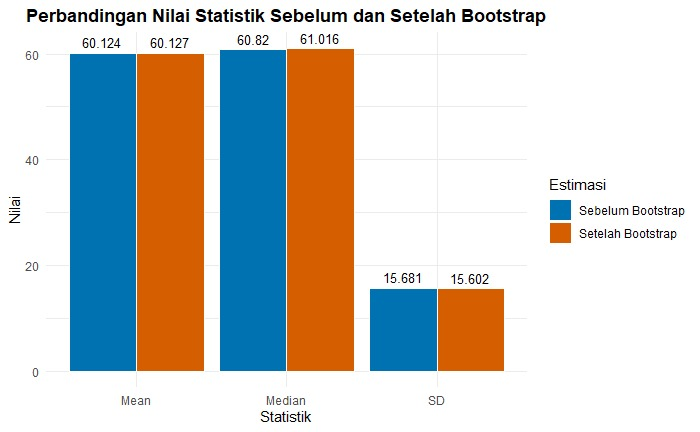
Distribusi *median* bootstrap juga tampak mengelompok secara kuat di sekitar nilai 61. Distribusi ini cenderung lebih sempit dibandingkan distribusi *mean*, menunjukkan bahwa *median* merupakan estimasi yang lebih robust dan kurang dipengaruhi oleh variasi sampel. Puncak distribusi yang tajam menunjukkan konsistensi nilai *median* pada berbagai replikasi bootstrap.

Pada distribusi standar deviasi (SD), terlihat bahwa penyebaran nilai lebih lebar dibandingkan *mean* dan *median*. Hal ini wajar, mengingat SD lebih sensitif terhadap perubahan kecil pada data individual. Meskipun demikian, distribusi SD tetap menunjukkan pola yang mendekati normal, dengan pusat distribusi berada di sekitar nilai 16. Garis merah vertikal kembali menunjukkan bahwa nilai awal SD berada cukup dekat dengan pusat distribusi bootstrap, sehingga dapat disimpulkan bahwa variabilitas data IPLM 2024 tidak mengalami perubahan besar setelah dilakukan resampling.

Secara keseluruhan, visualisasi ini menunjukkan bahwa ketiga statistik deskriptif utama *mean*, *median*, dan standar deviasi memiliki kestabilan yang baik setelah dilakukan proses bootstrap. Distribusi yang menyerupai bentuk lonceng (bell-shaped) menambah keyakinan bahwa sampel yang digunakan cukup representatif. Hasil ini memperkuat kesimpulan bahwa estimasi awal yang dihitung dari data IPLM 2024 dapat dipercaya, serta confidence interval yang diperoleh dari bootstrap memiliki dasar distribusi yang kuat.

## Perbandingan Statistik Sebelum dan Setelah Bootstrap

Selain menampilkan distribusi hasil bootstrap, penelitian ini juga menyajikan perbandingan langsung antara nilai statistik deskriptif sebelum dan setelah proses bootstrap. Visualisasi ini ditampilkan dalam bentuk diagram batang (bar chart) untuk tiga parameter utama, yaitu *mean*, *median*, dan standar deviasi (SD). Perbandingan ini bertujuan untuk melihat sejauh mana proses resampling mempengaruhi nilai estimasi yang diperoleh dari data IPLM 2024.



Gambar 4.5 Perbandingan Nilai Statistik Sebelum dan Sesudah Bootstrap

Berdasarkan Gambar 4.5 menunjukkan bahwa nilai *mean* sebelum dan sesudah bootstrap berada pada kisaran yang hampir sama, yaitu sekitar 60. Hal ini mengindikasikan bahwa estimasi rata-rata IPLM relatif stabil meskipun dilakukan resampling sebanyak 5.000 kali. Dengan kata lain, data IPLM tidak menunjukkan variabilitas ekstrem yang dapat menggeser estimasi *mean* secara signifikan.

Nilai *median* sebelum dan setelah bootstrap juga menunjukkan pola yang konsisten, dengan kedua nilai berada di sekitar 61. Konsistensi ini memperkuat karakter *median* sebagai ukuran pemusatan yang robust, yang tetap stabil meskipun dilakukan pengambilan sampel ulang secara intensif. Hasil ini sejalan dengan teori statistika bahwa *median* cenderung tidak sensitif terhadap nilai ekstrim, sehingga perubahan pada sampel pun tidak menghasilkan perbedaan besar pada estimasinya.

Untuk standar deviasi, nilai sebelum dan setelah bootstrap juga menunjukkan kesamaan yang kuat, yakni berada di sekitar 16. Hal ini menunjukkan bahwa tingkat variabilitas antar kota dan kabupaten di Sumatera dalam capaian IPLM 2024 relatif terjaga stabil, dan bootstrap tidak memberikan perubahan signifikan terhadap estimasi penyebaran data. Dengan demikian, dapat disimpulkan bahwa distribusi nilai IPLM pada kota dan kabupaten di Pulau Sumatera tidak mengalami fluktuasi besar yang dapat mengganggu estimasi SD. Secara keseluruhan, grafik perbandingan ini mempertegas bahwa proses bootstrap tidak hanya menghasilkan distribusi statistik yang stabil, tetapi juga memberikan bukti empiris bahwa nilai estimasi awal (observed statistics) sudah cukup representatif. Kesamaan antara nilai sebelum dan sesudah bootstrap menunjukkan bahwa sampel IPLM 2024 memiliki reliabilitas tinggi, serta confidence interval yang dihasilkan melalui bootstrap dapat dijadikan dasar inferensi statistik yang kuat.

## 4.7. Diskusi Temuan Penelitian

Penelitian ini menerapkan metode bootstrap untuk mengevaluasi stabilitas estimasi statistik deskriptif Indeks Pembangunan Literasi Masyarakat (IPLM) 2024. Temuan yang diperoleh menunjukkan konsistensi dengan teori serta temuan dari studi sebelumnya. Studi tersebut menjelaskan bahwa semakin besar ukuran resampling (B), nilai *Mean square error* (MSE) semakin menurun sehingga estimator menjadi semakin tidak bias. Konsistensi ini terlihat pada hasil estimasi kurtosis dan skewness, di mana resampling sebanyak 50, 100, 200, hingga 1000 menghasilkan nilai MSE terkecil serta varians yang mendekati MSE, sehingga estimator dapat dinyatakan tak bias [6].

Hasil penelitian ini memperkuat temuan tersebut. Meskipun penelitian ini tidak mengevaluasi kurtosis maupun skewness, bootstrap diaplikasikan untuk mengestimasi stabilitas *mean*, *median*, dan standar deviasi. Setelah dilakukan resampling sebanyak 5.000 kali, nilai estimasi *mean* (60.12722), *median* (61.01635), dan standar deviasi (15.60152) menunjukkan perubahan yang sangat kecil dibandingkan data asli. Dengan kata lain, bootstrap menghasilkan distribusi statistik yang lebih sempit dan terpusat, sehingga estimasi parameter menjadi lebih stabil. Temuan ini sejalan dengan pernyataan pada jurnal bahwa semakin besar jumlah sampel bootstrap, maka ketidakpastian estimasi semakin menurun, serta nilai estimasi semakin mendekati nilai parameter populasi sebenarnya.

# BAB V

**KESIMPULAN DAN SARAN**

## Kesimpulan

Berdasarkan hasil analisis terhadap data Indeks Pembangunan Literasi Masyarakat (IPLM) 2024 dan penerapan metode bootstrap sebanyak 5.000 replikasi, diperoleh beberapa temuan penting. Kesimpulan ini dirumuskan untuk menjawab tujuan penelitian dan menggambarkan efektivitas penggunaan bootstrap dalam proses estimasi statistik.

1. Metode bootstrap mampu menghasilkan estimasi statistik IPLM yang lebih stabil dibandingkan estimasi langsung tanpa resampling. Hal ini terlihat dari konsistensi nilai *mean* (sebelum bootstrap = 60.12356, setelah bootstrap = 60.12722), *median* (sebelum = 60.82, setelah bootstrap = 61.01635), dan standar deviasi (sebelum = 15.68139, setelah bootstrap = 15.60152). Perbedaan yang sangat kecil menunjukkan bahwa bootstrap memperkuat reliabilitas estimasi tanpa mengubah interpretasi nilai capaian IPLM secara substantif.
2. Distribusi IPLM masyarakat tahun 2024 berhasil digambarkan secara komprehensif melalui visualisasi histogram dan interval kepercayaan berbasis bootstrap. Interval kepercayaan 95% untuk *mean* (57.47392 – 62.86831), *median* (58.41 – 63.38), dan standar deviasi (13.94621 – 17.28635) menunjukkan rentang parameter populasi yang realistis dan dapat dipercaya, sehingga estimasi tidak bergantung semata-mata pada sampel tunggal.
3. Hasil penelitian menunjukkan bahwa metode bootstrap dapat menjadi pendekatan analitis yang mendukung proses perencanaan dan pengembangan program literasi berbasis data. Estimasi yang stabil memungkinkan identifikasi wilayah yang membutuhkan intervensi secara lebih tepat, karena variasi antar daerah tetap terbaca melalui standar deviasi dan confidence interval yang diperoleh.

## Saran

Sebagai tindak lanjut dari temuan penelitian, beberapa rekomendasi yang dapat dipertimbangkan untuk pengembangan penelitian dan implementasi kebijakan adalah sebagai berikut:

1. Penerapan teknik bootstrap disarankan untuk penelitian literasi berikutnya dengan cakupan wilayah nasional atau multiyear, sehingga pola perubahan literasi dapat dipetakan secara lebih dinamis dari tahun ke tahun.
2. Pemerintah daerah dan lembaga penggerak literasi diharapkan memanfaatkan estimasi statistik berbasis bootstrap sebagai dasar penyusunan program, terutama untuk menentukan prioritas daerah intervensi berdasarkan sebaran nilai IPLM dan tingkat ketimpangan antar wilayah.
3. Penelitian lanjutan dapat mengkombinasikan bootstrap dengan metode statistik lain seperti regresi, pemodelan prediksi, atau clustering guna memperoleh analisis yang lebih mendalam mengenai faktor-faktor yang mempengaruhi capaian literasi di suatu daerah.

# DAFTAR PUSTAKA

[1] G. Chairunisa, M. K. Najib, S. Nurdiati, S. F. Imni, W. Sanjaya, R. D. Andriani, … and D.  Ekaputri, “Life expectancy prediction using decision tree, random forest, gradient boosting, and XGBoost regressions,” *Jurnal Sintak*, vol. 2, no. 2, pp. 71–82, 2024.

[2] R. Kusumawati, “The strategi pengembangan desa digital untuk meningkatkan literasi, pelayanan publik, dan kemandirian masyarakat di Jawa Barat,” *Jurnal Governansi*, vol. 10, no. 1, pp. 25–40, 2024.

[3] A. Susilawati, M. Mursyid, and A. I. Madani, “Pengaruh religiusitas, literasi, dan akuntabilitas terhadap minat berwakaf tunai di Bank Syariah Indonesia dengan tingkat pemahaman sebagai variabel moderasi,” *Jurnal Syntax Admiration*, vol. 5, no. 10, pp. 4024–4037, 2024.

[4] N. Nurfajriyani, D. D. Amaliana, and S. P. Wulandari, “Pengelompokan provinsi berdasarkan aspek pembangunan pendidikan di Indonesia tahun 2023 menggunakan analisis cluster,” *Pentagon: Jurnal Matematika dan Ilmu Pengetahuan Alam*, vol. 2, no. 4, pp. 134–153, 2024.

[5] Eliani, Rais, and Fadjryani, “Analisis pengelompokan derajat kesehatan ibu dan anak di Indonesia menggunakan Structural Equation Modeling Partial Least Square-Prediction Oriented Segmentation (SEM PLS-POS),” *Jurnal Matematika, Statistika dan Komputasi*, vol. 19, no. 3, pp. 473–483, 2023.

[6] H. Sihombing and H. Simamora, “Metode Jackknife dan Metode Bootstrap dalam Estimasi Kurtosis dan Skewness,” *KARISMATIKA: Kumpulan Artikel Ilmiah, Informatika, Statistik, Matematika dan Aplikasi*, vol. 8, no. 2, pp. 144–156, 2022.

[7] C. E. Nababan and E. Simamora, “Bootstrap Estimation of Confidence Intervals of Multiple Regression Model Parameters in the Presence of Multicollinearity Using Principal Component Analysis,” *Formosa Journal of Applied Sciences*, vol. 2, no. 1, pp. 185–202, Jan. 2023, doi: 10.55927/fjas. v2i1.2851.

[8] A. S. Suherman, “The Bootstrap Percentile Confidence Interval untuk Koefisien Variasi,” *Jurnal Riset Statistika*, vol. 4, no. 2, pp. 109–116, 2024, doi: 10.29313/jrs.v4i2.3854.

[9] J. O. Fallo, A. Setiawan, and B. Sussanto, “Analisis kebutuhan pangan pokok pada provinsi-provinsi di Indonesia menggunakan indeks Moran berdasarkan metode bootstrap,” *Jurnal Sains Matematika dan Statistika*, vol. 6, no. 2, pp. 68–78, 2020.

[10] A. Fernandes, “Comparison of parameter estimator efficiency levels of path analysis with bootstrap and jack knife (delete-5) resampling methods on simulation data,” *Jurnal Matematika, Statistika dan Komputasi*, vol. 16, no. 3, pp. 353–364, 2020.

[11] T. W. Abadi and H. Sukmana, *Buku Ajar Statistika Deskriptif untuk Riset Komunikasi*.   Umsida Press, 2022.

[12] S. B. Pasaribu, *Statistika untuk Ekonomi dan Bisnis*. Universitas Indonesia, 2021.

[13] N. W. Sunita and A. Syakur, “Penerapan Statistik Deskriptif: Perspektif Kuantitatif dan Kualitatif,” *Emasains: Jurnal Edukasi Matematika dan Sains*, vol. 14, no. 1, pp. 1–15, 2025.

[14] N. Sa’adah, “Analisis Deskriptif Distribusi dan Tren Nilai Mahasiswa Mata Kuliah Kalkulus Periode 2019–2023,” *SUDO: Jurnal Teknik Informatika*, vol. 4, no. 3, 2025, doi: 10.56211/sudo. v4i3.1233.

[15] E. Roflin, Rohana, and F. Riana, *Statistika Dasar*. NoerFikri Offset, 2020.

[16] Badan Pusat Statistik, *Indeks Pembangunan Literasi Masyarakat dan Unsur-unsurnya Menurut Provinsi, 2024*. Jakarta: Badan Pusat Statistik, 2024.

[17] Perpustakaan Nasional Republik Indonesia, “IPLM 2024 Catat Rekor Tinggi, Literasi Nasional Semakin Meningkat,” 2024.

[18] Dispusip Sumut, *Capaian Indeks Pembangunan Literasi Masyarakat Provinsi Sumatera Utara Tahun 2023*. Dinas Perpustakaan dan Kearsipan Provinsi Sumatera Utara, 2024.