

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 *Big Data*

Fenomena *Big Data* muncul bersamaan dengan bertambahnya aktivitas digital di kalangan masyarakat, di mana hampir setiap interaksi, transaksi, dan proses operasional menghasilkan sejumlah besar jejak data (Fajriyah, dkk., 2022). *Big Data* biasanya dicirikan oleh lima unsur utama (5V), yaitu *volume*, *velocity*, *variety*, *veracity*, dan *value*. *Volume* mengacu pada ukuran data yang sangat besar, *velocity* menggambarkan kecepatan pengumpulan data secara berkelanjutan, *variety* menunjukkan berbagai format dan sumber data, *veracity* berkaitan dengan ketidakpastian kualitas data, sedangkan *value* menekankan kemampuan untuk mengambil nilai informasi dari data tersebut.

Untuk mengatur dan menganalisis data dengan karakteristik tersebut, diperlukan sebuah platform analitik dan infrastruktur komputasi yang dapat memproses data secara efektif, baik dari segi performa maupun efisiensi penggunaan sumber daya, untuk tujuan deskriptif, prediktif, maupun preskriptif (Hindayani, dkk., 2022).

2.1.2 *Big Data Analytics*

Big Data Analytics adalah teknologi yang memungkinkan pengolahan dan analisis sekumpulan data dalam jumlah besar dari berbagai sumber untuk memperoleh wawasan yang penting dan mendukung pengambilan keputusan (Gandomi, dkk., 2023). Teknologi ini tidak hanya menyediakan akses cepat ke basis

data besar, tetapi kemungkinan pengolahan data dalam skala yang sangat luas untuk menemukan wawasan tersembunyi yang dapat dimanfaatkan dalam strategi bisnis dan penelitian.

Proses analisis ini menggunakan algoritma maju dan model pembelajaran mesin untuk menemukan pola, tren, dan hubungan yang tidak terlihat dalam analisis konvensional. Namun, penggunaan *Big Data Analytics* menghadapi tantangan khusus, terutama terkait dengan kebutuhan daya komputasi yang tinggi. Pendekatan *Green Computing* menjadi perspektif penting untuk memastikan bahwa proses analisis data yang memerlukan daya pemrosesan tinggi tetap dapat dilaksanakan dengan efisiensi energi yang optimal (Gandomi, dkk., 2023). Dalam penelitian ini, istilah *Big Data Analytics* digunakan untuk merujuk pada pendekatan *pipeline* analitik, sementara pengujian dilakukan pada analitik data skala menengah pada perangkat kelas menengah sesuai rancangan eksperimen.

2.1.3 *Green Computing*

Green Computing adalah proses pengembangan dan optimalisasi *chip*, sistem, jaringan, dan perangkat lunak komputer dengan tujuan untuk memaksimalkan efisiensi energi dan meminimalkan dampak negatif terhadap lingkungan (Zhou, dkk., 2023). Seiring dengan meningkatnya ketergantungan pada teknologi informasi, penerapan *Green Computing* semakin krusial dalam menghasilkan ekosistem digital yang lebih ramah lingkungan serta berkelanjutan (Paul, dkk., 2023).

Dalam sistem analitik berskala menengah, seperti yang digunakan dalam *Big Data Analytics*, pengelolaan energi, *CPU*, dan memori menjadi faktor penting

yang turut menentukan performa suatu platform. Prinsip *Green Computing* dapat dimanfaatkan sebagai kerangka evaluatif untuk menilai seberapa efisien suatu platform menjalankan proses analitik, baik dari sisi pemanfaatan sumber daya komputasi maupun dampak lingkungan.

Green software menekankan hubungan antara performa komputasi, penggunaan sumber daya perangkat keras, dan konsumsi energi selama proses eksekusi. *Green Software Foundation* (GSF) menetapkan prinsip *energy efficiency*, *hardware efficiency*, dan *resource utilization* sebagai parameter dalam mengevaluasi perangkat lunak berkelanjutan. Selain itu, pengukuran *green software* juga dapat dilakukan melalui penggunaan sumber daya komputasi seperti penggunaan *CPU*, memori, penyimpanan, dan jaringan untuk merepresentasikan hubungan antara performa komputasi dan konsumsi energi perangkat lunak. Pendekatan ini menunjukkan bahwa parameter waktu eksekusi, utilisasi *CPU*, penggunaan memori, dan konsumsi energi dapat digunakan sebagai representasi performa berbasis *Green Computing*. Namun, parameter lain seperti *carbon intensity* dan *network efficiency* tidak menjadi fokus dalam penelitian ini (Brunnert, 2024; Green Software Foundation, 2025)

Dalam konteks penelitian ini, prinsip *Green Computing* digunakan sebagai pendekatan evaluatif untuk menilai performa platform *R* (melalui *RStudio*) dan *Python* secara menyeluruh, dengan mengukur konsumsi energi, penggunaan prosesor, pemanfaatan memori, dan waktu eksekusi sebagai indikator utama. Penelitian ini mengevaluasi performa platform berdasarkan parameter utama *Green Computing* berikut:

Tabel 2.1 Parameter Utama *Green Computing*

No.	Parameter	Indikator Pengukuran	Satuan	Interpretasi/ Rujukan
1.	Waktu Eksekusi	Total waktu yang dibutuhkan untuk menyelesaikan analisis.	Milidetik (ms)	Lebih rendah → performa lebih baik pada <i>workload</i> setara; konsistensi dilihat dari <i>mean/peak/P95</i> .
2.	Utilisasi <i>CPU</i>	Rata-rata dan puncak utilisasi <i>CPU</i> selama eksekusi.	Persentase (%)	Lebih rendah → pemanfaatan <i>CPU</i> lebih efisien pada <i>workload</i> setara; konsistensi dilihat dari <i>mean/peak/P95</i> .
3.	Penggunaan Memori	Rata-rata dan puncak penggunaan memori selama proses.	<i>Megabyte/Gigabyte</i> (MB/GB)	Lebih rendah → pemanfaatan memori lebih efisien pada <i>workload</i> setara; konsistensi dilihat dari <i>mean/peak/P95</i> .
4.	Konsumsi Energi	Total energi selama proses analitik.	J (<i>Joule</i>)	Lebih rendah → lebih efisien energi pada <i>workload</i> setara;

				konsistensi dilihat dari <i>mean/peak/P95.</i>
--	--	--	--	---

2.1.4 Performa Sistem Komputasi

Performa sistem komputasi menjadi salah satu aspek sentral dalam mengevaluasi keandalan dan efisiensi platform analitik, khususnya dalam pengolahan data pada *pipeline* analitik. Performa sistem komputasi merupakan ukuran kemampuan sistem dalam menjalankan tugas-tugas komputasi dengan efisien dan efektif. Evaluasi performa menjadi aspek krusial yang menentukan kualitas dan kelayakan suatu platform analitik (Mahmoudian, dkk., 2023).

Evaluasi performa sistem komputasi dapat dilakukan melalui berbagai dimensi yang saling berkaitan. Menurut (Li, dkk., 2023), dimensi utama meliputi *throughput* yang mengukur jumlah operasi yang dapat diselesaikan dalam satuan waktu tertentu, *latency* yang menunjukkan waktu respons sistem terhadap suatu permintaan, dan *scalability* yang menggambarkan kemampuan sistem untuk mempertahankan performa saat beban kerja meningkat. Sebagai tambahan, *resource utilization* menjadi indikator penting yang mengukur efisiensi penggunaan sumber daya seperti *CPU*, memori, dan *storage*.

Dalam era *Big Data*, karakteristik khusus dari volume, kecepatan, dan variasi data menimbulkan tantangan tersendiri bagi evaluasi performa sistem. Studi terbaru mengindikasikan bahwa evaluasi performa sistem komputasi dalam *Big Data Analytics* memerlukan pendekatan multidimensi yang mengintegrasikan aspek teknis dan operasional. Menurut (Aljehani, dkk., 2024), metrik performa

modern harus mencakup *execution time* untuk mengukur kecepatan pemrosesan, *memory efficiency* untuk mengevaluasi optimalisasi penggunaan memori, dan *CPU utilization* untuk menganalisis beban kerja prosesor.

Penelitian oleh (Theodorakopoulos, dkk., 2024), Menekankan pentingnya metodologi *benchmarking* yang standar dalam evaluasi performa. Metodologi *benchmark* yang tepat tidak hanya memberikan hasil yang dapat direproduksi, tetapi juga memungkinkan perbandingan yang objektif antara berbagai platform analitik. Dalam konteks *sustainability* dan *Green Computing*, performa sistem komputasi tidak lagi hanya diukur dari aspek kecepatan dan efisiensi, tetapi juga dari perspektif konsumsi energi dan dampak lingkungan. (Aljehani, dkk., 2024) menunjukkan bahwa integrasi parameter *Green Computing* dalam evaluasi performa memberikan perspektif yang lebih komprehensif tentang kualitas sistem komputasi dalam jangka panjang. Pemahaman dimensi-dimensi performa ini memberikan dasar evaluasi yang kuat dalam membandingkan efektivitas *R* (melalui *RStudio*) dan *Python* sebagai platform analitik berbasis *Green Computing*.

2.1.5 Python

Python merupakan bahasa pemrograman berorientasi objek yang dapat digunakan untuk pengembangan perangkat lunak serta kompatibel dengan berbagai sistem operasi. Dalam *Big Data Analytics*, *Python* menjadi salah satu platform yang paling banyak digunakan karena menyediakan berbagai pustaka yang mendukung analisis data, *machine learning*, *preprocessing* data, serta visualisasi data (Manalu dan Gunadi, 2022).

Keunggulan utama *Python* dalam *Big Data Analytics* terletak pada fleksibilitas dan ekosistem pustaka yang luas, seperti *pandas*, *numpy*, dan *scikit-learn*, yang mempermudah pengolahan serta analisis data dalam skala besar. Sebagai tambahan, *Python* mendukung integrasi dengan berbagai teknologi modern seperti *cloud computing* dan *distributed computing*, sehingga semakin memperkuat posisinya sebagai bahasa yang efisien untuk mengolah *Big Data*.

Dalam konteks *Green Computing*, efisiensi *Python* dalam penggunaan sumber daya menjadi faktor penting yang perlu dievaluasi. Dari hal tersebut, penting dilakukan kajian komparatif yang menilai performa *Python* secara menyeluruh dalam konteks *Green Computing*, agar dapat dibandingkan secara objektif dengan platform lain seperti *R* (melalui *RStudio*).

2.1.6 Bahasa Pemrograman *R* dan *RStudio*

R merupakan salah satu bahasa pemrograman yang dibuat khusus untuk analisis statistik, komputasi ilmiah, dan visualisasi data. Ekosistem *R* didukung oleh sejumlah paket yang menyediakan fungsional untuk pemodelan statistik, analisis deret waktu, pengujian hipotesis, serta pembelajaran mesin yang telah diterapkan secara luas di berbagai sektor, termasuk keuangan dan penilaian risiko (Savita dan Verma, 2020).

RStudio adalah *Integrated Development Environment (IDE)* yang dirancang untuk menyederhanakan penulisan, eksekusi, dan pengelolaan skrip *R* melalui antarmuka yang teratur. Dalam konteks *Big Data Analytics*, *R* yang dioperasikan melalui *RStudio* dapat digunakan untuk melakukan eksplorasi data, transformasi, dan visualisasi dengan cara efisien, terutama bila dipadukan dengan paket-paket

yang meningkatkan pengelolaan data dalam skala besar. Penelitian ini menganalisis performa *R* (melalui *RStudio*) dengan membandingkan terhadap *Python* dari perspektif *Green Computing*, yakni dengan mengkaji parameter waktu eksekusi, utilisasi *CPU*, penggunaan memori, dan konsumsi energi.

2.2 Penelitian Terkait dan Kebaruan Penelitian

2.2.1 Penelitian Terkait

Komparasi performa antara *R* (melalui *RStudio*) dan *Python* dalam konteks *Big Data Analytics* telah menjadi topik yang berkembang dalam beberapa tahun terakhir. Berbagai penelitian telah mengevaluasi kedua platform dari perspektif beragam, mulai dari waktu eksekusi, utilisasi sumber daya komputasi, kemampuan visualisasi, hingga interoperabilitas antar platform. Namun, evaluasi yang mengintegrasikan aspek konsumsi energi sebagai parameter performa dalam kerangka *Green Computing* masih terbatas. Tinjauan ini merangkum penelitian terdahulu yang relevan sebagai landasan konseptual dan empiris bagi penelitian ini.

Dari perspektif komparasi teknis, (Dipto, dkk., 2020) menguji kemampuan analitik *Python* dan *R* menggunakan *Artificial Neural Network* (ANN) untuk memprediksi tingkat keparahan kecelakaan. Studi tersebut menggabungkan dua dataset besar dari *Kaggle* yang diolah melalui tahapan visualisasi data dan *preprocessing* untuk pelatihan ANN dengan algoritma *Multilayer Perceptron*. Hasil penelitian menyoroti kelebihan dan kekurangan setiap bahasa dalam hal efisiensi *preprocessing* data, kemampuan visualisasi, dan akurasi hasil prediktif, memberikan wawasan penting tentang karakteristik kedua platform dalam analitik data skala besar. Pemahaman mengenai interoperabilitas kedua bahasa diperluas

oleh (Scavetta, 2022) yang menyusun panduan komprehensif mengenai penggunaan *Python* dan *R* untuk mendukung berbagai tantangan statistik modern, seperti analisis data eksploratif, pembelajaran mesin, rekayasa data, dan pelaporan. Panduan ini terstruktur dalam tujuh bab yang mencakup sejarah, konteks, perbandingan kedua bahasa, serta menyediakan contoh kerja dan sumber daya tambahan untuk mendukung pengambilan keputusan berbasis data. Selain itu, (Hulsen, 2021) mengembangkan paket lintas *R* dan *Python* untuk visualisasi dan perbandingan daftar biologis menggunakan diagram Venn proporsional. Studi ini menunjukkan bahwa kedua ekosistem mampu mendukung pengembangan alat analitik dan visualisasi data secara fleksibel, sehingga relevan sebagai bukti kuatnya dukungan pustaka pada masing-masing platform. Penelitian ini menunjukkan tingkat fleksibilitas tinggi kedua bahasa dalam pengembangan alat visualisasi data khusus.

Dalam konteks performa dan konsumsi energi, fokus penelitian (Nahrstedt, dkk., 2024) melakukan studi empiris mengenai penggunaan energi dan performa pustaka analisis data *Python*, khususnya *pandas* dan *polars*. Penelitian ini melibatkan empat eksperimen yang mencakup 8 *Data Analysis Tasks* (DATs) dari *TPCH Benchmark* dan 6 *DAT* sintetik yang diuji pada *dataframe* berukuran kecil dan besar. Eksperimen dilakukan menggunakan alat seperti *PowerJoular* untuk mengukur konsumsi energi dan *CPU*, serta *PS* untuk penggunaan memori. Hasil penelitian menunjukkan bahwa *polars* memiliki performa lebih baik dengan konsumsi energi yang lebih rendah dibandingkan *pandas*, terutama untuk pengolahan data berukuran besar, serta memberikan panduan bagi ilmuwan data

dalam memilih pustaka berdasarkan kebutuhan performa dan konsumsi energi. Relevansi aspek konsumsi energi ini sejalan dengan konsep *Green Computing* yang dijelaskan oleh (Sneha, dkk., 2021), yang mengeksplorasi penerapan *Big Data Analytics* dan pembelajaran mesin dalam mendukung komputasi hijau dengan fokus pada efisiensi energi dan pengurangan dampak lingkungan dari teknologi informasi. Metode penelitian melibatkan analisis data besar yang dihasilkan oleh manusia, mesin, organisasi, serta penerapan algoritma pembelajaran mesin seperti pembelajaran terawasi, tidak terawasi, dan penguatan. Penelitian ini mencakup tahapan siklus hidup *Big Data*, mulai dari *Spark* dan *NoSQL* untuk memastikan efisiensi dan skalabilitas dalam solusi komputasi berkelanjutan.

Implementasi *Big Data Analytics* (BDA) juga telah diterapkan secara luas di berbagai sektor. Dalam domain kesehatan, (Batko dan Ślęzak, 2022) menganalisis potensi penggunaan BDA di fasilitas medis Polandia melalui kombinasi analisis kritis literatur dengan penelitian langsung menggunakan kuesioner yang dikirimkan kepada 217 fasilitas medis. Penelitian ini menunjukkan bahwa fasilitas kesehatan secara bertahap beralih ke sistem berbasis data, memanfaatkan teknologi BDA untuk mendukung pengambilan keputusan berbasis data guna meningkatkan kualitas layanan kesehatan, memahami kebutuhan pasien, dan memberikan perawatan yang lebih personal. Sejalan dengan hal tersebut, (Gates, dkk., 2024) memanfaatkan analitik data besar untuk mengidentifikasi pola prediktif yang dapat meningkatkan efisiensi pelayanan kesehatan. Penelitian ini melibatkan analisis data historis dan prospektif menggunakan teknologi pengolahan data, algoritma pembelajaran mesin, dan teknik visualisasi data untuk

mengidentifikasi kesehatan serta memberikan rekomendasi berbasis data di fasilitas medis. Secara lebih spesifik, (Gandhi, dkk., 2024) memperkenalkan perangkat lunak *R* sebagai alat analisis data yang efektif dalam penelitian kanker. Penelitian ini menyediakan panduan data *dummy*, penjelasan terperinci, serta *screenshot* kode untuk membantu pengguna non-programmer memahami analisis data medis menggunakan *R*, sehingga dapat berkontribusi secara mandiri pada penelitian kanker dengan fleksibilitas dan akurasi yang lebih baik.

Penerapan BDA juga merambah ke sektor publik dan industri, pada penelitian (Dharma dan Hendri, 2022) mengkaji penerapan BDA pada audit sektor publik melalui analisis SWOT dan menyimpulkan bahwa BDA dapat meningkatkan efisiensi serta efektivitas audit, terutama melalui pemanfaatan data terstruktur dan tidak terstruktur. Dalam konteks kota pintar, (Talebkhah, dkk., 2023) mengkaji integrasi teknologi industri 4.0 terutama *Big Data*, *Internet of Things* (IoT) dan *cloud computing* dalam pengembangan kota pintar, serta menyoroti tantangan teknis dan kebutuhan arsitektur hibrida untuk mendukung perencanaan dan pengelolaan perkotaan berbasis sensor. Pada ranah bisnis, (Mahananing Puri, dkk., 2024) memanfaatkan *Big Data* dari *Twitter* untuk rekomendasi produk dengan analisis menggunakan *RStudio*. Data dikelompokkan melalui *K-Means Clustering* dan diberi peringkat menggunakan metode *Simple Additive Weighting* (SAW), dengan evaluasi *clustering* menggunakan *Davies-Bouldin Index* (DBI). Sementara itu, (Tyagi, 2024) merangkum perkembangan teknik *Big Data Analytics* berbasis *machine learning*, termasuk *deep learning*, *neural networks*, dan *decision tree*, serta

menekankan perannya dalam menghadapi kompleksitas data dan kebutuhan pengambilan keputusan berbasis data.

Secara keseluruhan, penelitian terdahulu menunjukkan bahwa komparasi *R* dan *Python* dipengaruhi oleh karakteristik teknis serta pola penggunaan pada beban kerja tertentu (Dipto, dkk., 2020; Scavetta, 2022). Efektivitas implementasi bergantung pada ketersediaan alat bantu dan fleksibilitas pengembangan perangkat lunak analitik dalam ekosistem (Hulsen, 2021). Sejalan dengan meningkatnya ukuran data dan intensitas komputasi, efisiensi energi dalam kerangka *Green Computing* mulai dipertimbangkan sebagai parameter penting dalam pemilihan teknologi analitik data (Nahrstedt, dkk., 2024; Sneha, dkk., 2021). Dari sisi penerapan, bukti lintas sektor kesehatan, publik, kota pintar, dan bisnis mengindikasikan bahwa pemilihan platform analitik perlu dipertimbangkan performa sekaligus efisiensi sumber daya agar implementasi *Big Data Analytics* dapat berlangsung secara berkelanjutan (Batko dan Ślęzak, 2022; Dharma dan Hendri, 2022; Gandhi, dkk., 2024; Gates, dkk., 2024; Mahananing Puri, dkk., 2024; Talebkhah, dkk., 2023; Tyagi, 2024). Dalam kajian yang mengintegrasikan evaluasi konsumsi energi dalam komparasi lintas ekosistem *R* dan *Python* dengan beban kerja yang disetarakan masih terbatas, sehingga diperlukan pengujian empiris yang lebih terukur untuk memperkuat dasar pemilihan platform analitik.

2.2.2 Matriks Penelitian

Matriks penelitian disusun untuk memetakan fokus penelitian terdahulu dan menegaskan posisi penelitian yang dilakukan.

Tabel 2.2 Matriks Penelitian

No.	Penulis	Pemrograman				Software		Metode Pengembangan					Metode Pengujian		
		<i>Python</i>	<i>Hadoop</i>	<i>Spark</i>	<i>NoSQL</i>	<i>RStudio</i>	<i>GIT Solution</i>	<i>Green Computing</i>	<i>SWOT</i>	<i>K-Means</i>	<i>Data Analytics</i>	<i>ANN</i>	<i>SAW</i>	<i>DBI</i>	Kompatibilitas & Lainnya
1.	<i>Prediction of Accident Severity Using Artificial Neural Network: A Comparison of Analytical Capabilities between Python and R, (Dipto, dkk., 2020).</i>	✓				✓						✓			✓
2.	<i>BioVenn– an R and Python package for the comparison and visualization of biological lists using area-proportional Venn diagrams, (Hulsen, 2021).</i>	✓				✓									✓
3.	<i>Big Data Analysis and Machine Learning for Green Computing, (Sneha, dkk., 2021).</i>	✓	✓	✓	✓	✓		✓			✓				✓
4.	<i>The use of Big Data Analytics in healthcare, (Batko dan Ślęzak, 2022).</i>										✓				✓

No.	Penulis	Pemrograman				Software		Metode Pengembangan					Metode Pengujian		
		<i>Python</i>	<i>Hadoop</i>	<i>Spark</i>	<i>NoSQL</i>	<i>RStudio</i>	<i>GIT Solution</i>	<i>Green Computing</i>	<i>SWOT</i>	<i>K-Means</i>	<i>Data Analytics</i>	<i>ANN</i>	<i>SAW</i>	<i>DBI</i>	<i>Kompatibilitas & Lainnya</i>
5.	<i>Python and R for the Modern Data Scientist</i> , (Scavetta, 2022).	✓				✓									✓
6.	Urgensi Penggunaan <i>Big Data Analytics</i> dalam Audit Sektor Publik, (Dharma dan Hendri, 2022).								✓		✓				✓
7.	<i>Comprehensive Review on Development of Smart Cities Using Industry 4.0 Technologies</i> , (Talebkhah, dkk., 2023).		✓			✓					✓				✓
8.	<i>Efficient Data Processing: Assessing the Performance of Different Programming Languages</i> , (Beierlieb, dkk., 2023).	✓				✓					✓				✓
9.	Analisis Rekomendasi Pembuatan Produk Menggunakan <i>RStudio</i> dan Twitter (Studi Kasus: <i>Git Solution</i>), (Mahananing Puri, dkk., 2024).					✓	✓		✓				✓	✓	
10.	<i>An Empirical Study on the Energy Usage and Performance of pandas and polars Data Analysis Python Libraries</i> , (Nahrstedt, dkk., 2024).	✓									✓				✓

No.	Penulis	Pemrograman				Software		Metode Pengembangan					Metode Pengujian		
		<i>Python</i>	<i>Hadoop</i>	<i>Spark</i>	<i>NoSQL</i>	<i>RStudio</i>	<i>GIT Solution</i>	<i>Green Computing</i>	<i>SWOT</i>	<i>K-Means</i>	<i>Data Analytics</i>	<i>ANN</i>	<i>SAW</i>	<i>DBI</i>	<i>Kompatibilitas & Lainnya</i>
11.	<i>Latest Trending Techniques and ML Models for Big Data Analytics - A Review, (Tyagi, 2024).</i>		✓	✓		✓					✓				✓
12.	<i>A narrative review with a step-by-step guide to R software for clinicians: Navigating medical data analysis in cancer research, (Gandhi, dkk., 2024).</i>					✓									✓
13.	<i>Big Data Analytics for Predictive Insights in Healthcare, (Gates, dkk., 2024).</i>	✓									✓				
14.	<i>Who Wins the Race? (R Vs Python) – An Exploratory Study on Energy Consumption of Machine Learning Algorithms, (Chattaraj, dkk., 2025).</i>	✓				✓		✓			✓	✓			✓
15.	<i>Comparative analysis of large data processing in Apache Spark using Java, Python and Scala, (Borodii, dkk., 2025).</i>	✓		✓							✓				✓
16.	<i>(Penelitian ini, 2025).</i>	✓				✓		✓			✓				✓

*Keterangan: (✓) = Kategori pengembangan yang dilakukan

2.2.3 Kebaruan Penelitian

Berdasarkan tinjauan matriks penelitian yang telah dilakukan, penelitian ini memiliki keterkaitan dengan studi terdahulu seperti Monica Sneha dkk., Batko & Ślęzak, serta Maheshwari, yang membahas performa platform analitik data berbasis *Python* maupun *RStudio*. Sebagian besar penelitian tersebut fokus pada satu platform atau membandingkan pustaka tertentu dengan skenario yang spesifik, tanpa mengintegrasikan prinsip *Green Computing* secara eksplisit dalam evaluasi performanya. Penelitian ini menawarkan kebaruan melalui pendekatan komparatif yang sistematis antara *RStudio* dan *Python*, dengan parameter pengukuran yang distandarisasi dan disetarakan, mencakup konsumsi energi, utilisasi *CPU*, memori, dan waktu eksekusi. Penelitian ini menyajikan analisis komparatif berbasis ringkasan statistik pengukuran (nilai rata-rata, nilai puncak, dan *P95*) untuk memperkuat interpretasi perbedaan performa dan efisiensi sumber daya. Analisis ini diposisikan untuk menilai konsistensi perbedaan pada skenario uji yang sama, sesuai rancangan eksperimen. Pendekatan ini diharapkan memberikan kontribusi empiris yang lebih kuat dan relevan terhadap pengembangan *Big Data Analytics* yang efisien dan ramah lingkungan.