

Pengujian Wilcoxon Signed Rank Satu Sampel Terhadap Penggunaan Aplikasi AI-G dengan Perangkat Jamovi pada Prodi Informatika UNIBANG 2026

One-Sample Wilcoxon Signed-Rank Test on AI-G Application Usage Data Using Jamovi Software at UNIBANG Informatics Study Program (2026)

Ahmad Nur Taufiqurrahman

Program Studi Informatika, Fakultas Ilmu Kesehatan dan Sains, Universitas Bhakti Asih Tangerang
Jl. Raden Fatah No.62, Kota Tangerang, Banten, Indonesia
Email: taufiqurrahman@univbhaktiasih.ac.id

Abstrak – Kemajuan pesat teknologi kecerdasan buatan (AI) kini menjadi instrumen penting dalam mendukung kegiatan akademik mahasiswa. Penelitian ini bertujuan untuk mengukur dan membandingkan tingkat kemampuan mahasiswa Program Studi Informatika Universitas Bhakti Asih Tangerang (UNIBANG) dalam mengoperasikan AI-G dengan nilai acuan. Pertanyaan utama penelitian meliputi: (1) apakah distribusi data kemampuan mahasiswa memenuhi asumsi normalitas parametrik, dan (2) apakah terdapat perbedaan signifikan antara tingkat kemampuan mahasiswa dan nilai acuan yang telah ditetapkan? Penelitian kuantitatif ini dilaksanakan sejak bulan Januari hingga April 2026. Pengumpulan data dilakukan melalui instrumen kuesioner sederhana. Terdapat tiga pertanyaan utama yang perlu dijawab. Perangkat statistik yang digunakan adalah JAMOMI 2.6. Metode yang digunakan adalah uji satu sampel dengan nilai acuan 3. Analisis data diawali dengan uji asumsi normalitas. Hasil pengujian menunjukkan bahwa data tidak terdistribusi secara normal dengan p-value (0,001–0,044). Sehingga analisis satu sampel yang awalnya dirancang menggunakan uji t satu sampel (*one-sample t-test*) dialihkan ke metode statistik nonparametrik, yaitu *Wilcoxon Signed-Rank*. Hasil analisis inferensial menunjukkan nilai p-value (0,135–0,850) yang berarti tidak ada perbedaan signifikan antara nilai statistik dan nilai acuan, dengan kecenderungan median berada di bawah acuan. Kesimpulan ini menyatakan bahwa responden masih jarang menggunakan aplikasi AI-G untuk membantu pembelajaran. Hal ini dapat diasumsikan bahwa mereka masih lebih memilih menggunakan aplikasi pembuat salindia konvensional. Ketidakhiasaan penggunaan AI seperti aplikasi AI-G dapat mengurangi kemampuan responden untuk mencapai kesuksesan atau memasuki dunia industri.

Kata kunci: Nonparametrik; Literasi Digital; Teknologi Informasi; Pembelajaran; Kecerdasan Buatan

Abstract - The rapid advancement of artificial intelligence (AI) technology has become a vital instrument in supporting students' academic activities in higher education. This research aims to measure and evaluate the proficiency level of Informatics students at Universitas Bhakti Asih Tangerang (UNIBANG) regarding the operation of AI-G tools, compared against an established standard reference value. The core research questions are: (1) does the distribution of student proficiency data satisfy the assumption of parametric normality, and (2) is there a statistically significant difference between the students' measured proficiency levels and the predefined reference standard? This quantitative study was conducted between January and April 2026. Data collection was performed using a simple questionnaire comprising three primary questions. The data analysis utilized JAMOMI 2.6 software, employing a one-sample methodology with a reference value of 3. The initial analysis began with a normality test; results indicated that the data were not normally distributed, with p-values ranging from 0.001 to 0.044. Consequently, the analytical approach was shifted from the initially planned one-sample t-test to the nonparametric Wilcoxon Signed-Rank test. The inferential results show p-values between 0.135 and 0.850, indicating no significant statistical difference between the students' scores and the reference value. However, the median scores remained below the designated threshold. These findings suggest that respondents rarely utilize AI-G applications for academic support, likely due to a persistent preference for conventional presentation software. Unfamiliarity with AI-G applications may hinder students' ability to achieve academic success and reduce their overall competitiveness when entering the professional industry.

Keywords: Nonparametric; Digital Literacy; Information Technology; Learning; Artificial Intelligence

1. PENDAHULUAN

Melalui kemajuan teknologi kecerdasan buatan (*artificial intelligence/AI*) yang semakin pesat, institusi pendidikan semakin banyak menggunakan alat bantu berbasis AI (Dikkatwar *et al.*, 2025) seperti Meta, Gemini, Gamma, serta aplikasi lainnya untuk mendukung aktivitas akademik mahasiswa. Dalam konteks ini, penelitian mengenai tingkat kemampuan dan frekuensi penggunaan AI oleh mahasiswa menjadi penting sebagai dasar pengambilan keputusan dan perancangan pembelajaran. Pada tingkat metodologis, uji statistik seperti uji t satu sampel (*one-sample t-test*) merupakan salah satu perangkat yang penting untuk menilai apakah rata-rata suatu variabel populasi berbeda secara signifikan dari nilai acuan tertentu (Ross & Willson, 2017), misalnya: skor kemampuan atau frekuensi penggunaan AI. Uji t satu sampel digunakan untuk menguji hipotesis variabel kontinu terhadap nilai hipotesis tertentu, dengan mempertimbangkan variabilitas dan ukuran sampel.

Penelitian ini bertujuan untuk memetakan dan membandingkan tingkat kemampuan serta penggunaan aplikasi AI-G dengan nilai acuan, dalam aktivitas akademik mahasiswa Program Studi Informatika di Universitas Bhakti Asih Tangerang (UNIBANG) dengan menerapkan metode uji t satu sampel (jika data tidak normal, maka pengujian dilanjutkan dengan metode nonparametrik *Wilcoxon Signed-Rank* satu sampel). Secara khusus, penelitian ini berfokus pada (1) mengukur rata-rata kemampuan mahasiswa dalam memanfaatkan aplikasi AI-G untuk aktivitas akademik seperti pembuatan presentasi, penelitian, atau pemecahan masalah, dan (2) menguji apakah rata-rata tersebut secara statistik berbeda dari nilai acuan yang telah ditetapkan (misalnya, nilai standar minimal penggunaan AI). Dengan demikian, penelitian ini tidak hanya bersifat deskriptif, tetapi juga inferensial, memberikan dasar untuk pengambilan kesimpulan statistik yang lebih valid.

Meskipun telah banyak penelitian yang membahas penggunaan AI dalam pendidikan, terdapat celah riset (*research gap*) yang cukup jelas. Sebagian besar literatur sebelumnya fokus pada aspek persepsi, adopsi, atau hambatan penggunaan AI oleh dosen maupun mahasiswa (Rodzi *et al.*, 2023), namun sedikit yang mengukur secara kuantitatif menggunakan uji inferensial terhadap rata-rata kemampuan atau penggunaan mahasiswa dibandingkan dengan nilai acuan. Selain itu, penelitian yang mengaplikasikan uji t satu sampel/*Wilcoxon Signed-Rank* satu sampel dalam konteks penggunaan aplikasi AI-G untuk pendidikan, khususnya Prodi Informatika di Universitas Bhakti Asih Tangerang 2026, belum ditemukan dan merupakan sesuatu yang baru.

Penelitian ini bermanfaat untuk memberikan gambaran mengenai kemampuan mahasiswa dalam memanfaatkan AI, terutama aplikasi AI-G untuk kegiatan akademik. Temuan dari penelitian ini bisa menjadi landasan untuk mengembangkan strategi pembelajaran yang berbasis AI, memperkuat personalisasi dan kreativitas, bahkan merumuskan kebijakan institusional yang mendukung penggunaan AI secara transparan dan etis (Panagiotou, 2025), terutama penerapannya di Program Studi Informatika Fakultas Ilmu Kesehatan dan Sains UNIBANG. Oleh karena itu, penelitian ini memberikan kontribusi praktis dalam memperkuat kompetensi mahasiswa serta mengembangkan pendidikan tinggi yang responsif terhadap kemajuan teknologi AI.

Landasan Teori

Uji t satu sampel pertama kali diperkenalkan oleh William Sealy Gosset pada awal abad ke-20, khususnya pada tahun 1908, untuk mengatasi keterbatasan dalam pengujian statistik pada sampel kecil (Box, 1987). Teori ini berkembang sebagai metode inferensial yang memungkinkan peneliti menguji hipotesis mengenai rata-rata populasi dengan menggunakan data dari satu sampel tunggal. Konsep dasar uji t satu sampel adalah membandingkan rata-rata sampel dengan nilai rata-rata populasi yang diharapkan, dengan memperhitungkan variabilitas sampel dan ukuran sampel (Zabell, 2008). Seiring waktu, teori ini disempurnakan dan diperluas melalui kontribusi para ahli statistik, seperti Student (nama samaran Gosset) dan Ronald Fisher, yang mengembangkan distribusi t dan memperkenalkan pendekatan yang lebih sistematis dalam pengujian hipotesis (Rao, 1992; Noordzij *et al.*, 2010).

Perkembangan uji t satu sampel hingga tahun 2026 menunjukkan peningkatan aplikasi dan adaptasi metode ini dalam berbagai bidang penelitian, termasuk kesehatan dan teknik (Jankowski *et al.*, 2017; Deus *et al.*, 2024). Selain itu, kemajuan teknologi komputasi dan perangkat lunak statistik telah memungkinkan penerapan uji t satu sampel secara lebih efisien dan akurat. Penelitian terkini juga menyoroti pengembangan varian uji t yang lebih kuat terhadap pelanggaran asumsi klasik, seperti uji t dengan koreksi untuk data yang tidak berdistribusi normal atau memiliki varians yang tidak homogen (Mcgee, 2018). Dengan demikian, uji t satu sampel terus menjadi alat penting dalam analisis statistik yang mengalami evolusi baik dari segi teori maupun praktik hingga mencapai bentuk yang lebih adaptif dan relevan dengan kebutuhan penelitian modern. Perkembangan selanjutnya melibatkan perluasan jenis uji satu sampel untuk berbagai parameter, seperti median, proporsi, dan varians, serta adaptasi uji ini untuk data yang tidak memenuhi asumsi normalitas melalui uji nonparametrik, misalnya uji tanda (*sign test*) dan uji *Wilcoxon signed-rank* (Harris & Hardin, 2013; Nahm, 2016). Kemajuan dalam komputasi statistik mendorong penggunaan metode *bootstrap* dan simulasi Monte Carlo untuk meningkatkan akurasi estimasi dan pengujian hipotesis satu sampel, terutama dalam kondisi data yang kompleks atau berukuran kecil (Ichikawa & Konishi, 1995; Zhang, 2013).

Dalam era modern, integrasi uji statistik satu sampel dengan teknologi pembelajaran mesin dan kecerdasan buatan juga mulai muncul, terutama dalam konteks analisis data besar dan aplikasi pembelajaran adaptif (Destiani & Mufiidah, 2024). Penggunaan algoritma AI memungkinkan pengujian hipotesis yang lebih dinamis dan personalisasi analisis statistik berdasarkan karakteristik data yang lebih beragam. Hal ini membuka peluang baru dalam pengembangan metode pengajaran dan evaluasi pembelajaran berbasis aplikasi, seperti yang terlihat dalam penelitian penggunaan aplikasi Gamma untuk pembelajaran statistika. Pada praktiknya, uji statistik satu sampel tetap menjadi alat penting dalam penelitian pendidikan dan ilmu sosial untuk mengevaluasi efektivitas intervensi, termasuk penggunaan aplikasi pembelajaran digital. Penelitian terkini sering menggabungkan uji ini dengan analisis kualitatif dan model statistik lanjutan untuk memberikan pemahaman yang lebih menyeluruh tentang hasil dan dampak intervensi.

Perkembangan Terbaru Uji Satu Sampel

Tinjauan literatur terkini mengenai penggunaan aplikasi statistik Jamovi dalam pelaksanaan uji t satu sampel menunjukkan bahwa Jamovi semakin populer sebagai alat analisis statistik yang *user-friendly* dan efisien, terutama di kalangan akademisi dan peneliti yang membutuhkan proses analisis data yang cepat dan akurat tanpa harus menguasai pemrograman statistik yang kompleks (Navarro & Foxcroft, 2025). Jamovi menyediakan antarmuka grafis yang intuitif, memungkinkan pengguna untuk melakukan uji t satu sampel dengan mudah melalui beberapa langkah sederhana, sekaligus memberikan keluaran hasil yang lengkap berupa nilai t, derajat kebebasan, nilai p, interval kepercayaan, serta ukuran efek yang relevan untuk interpretasi hasil.

Beberapa studi terbaru menyoroti keunggulan Jamovi dalam meningkatkan *reproductibility* dan transparansi analisis statistik. Hal ini dikarenakan Jamovi mendukung ekspor hasil analisis dalam berbagai format yang mudah dibagikan dan dilaporkan, serta integrasi dengan R sebagai basis pemrogramannya memungkinkan fleksibilitas dalam pengembangan skrip analisis lanjutan (Şahin & Aybek, 2020). Penggunaan Jamovi dalam uji t satu sampel juga didukung oleh kemampuannya untuk menangani data dengan asumsi normalitas yang dapat diuji langsung di platform tersebut, sehingga peneliti dapat memastikan validitas uji sebelum melanjutkan ke interpretasi hasil.

Selain itu, literatur terkini menunjukkan bahwa Jamovi banyak digunakan dalam bidang-bidang seperti psikologi, pendidikan, dan informatika untuk menguji hipotesis yang melibatkan perbandingan rata-rata sampel dengan nilai populasi tertentu. Dalam konteks informatika, Jamovi membantu peneliti dalam evaluasi performa algoritma dan validasi eksperimen dengan mudah dan efisien (Abbasnasab Sardareh *et al.*, 2021). Kemudahan penggunaan dan fitur lengkap yang disediakan oleh Jamovi menjadikannya alat yang sangat mendukung proses analisis statistik yang akurat, transparan, dan mudah direplikasi oleh peneliti maupun mahasiswa.

Tren terkini dalam pengembangan aplikasi pembelajaran statistika, terutama uji t satu sampel, menunjukkan peningkatan pemanfaatan teknologi interaktif dan adaptif untuk meningkatkan efektivitas proses belajar, seperti uji t satu sampel berbasis web (Mustafidah *et al.*, 2020). Aplikasi-aplikasi modern kini mengintegrasikan fitur seperti visualisasi data dinamis, simulasi interaktif, dan pembelajaran berbasis masalah yang memungkinkan pengguna memahami konsep statistik secara lebih mendalam dan kontekstual. Selain itu, penggunaan kecerdasan buatan dalam aplikasi ini semakin berkembang, dengan kemampuan untuk menyesuaikan materi dan latihan berdasarkan tingkat pemahaman individu, sehingga memberikan pengalaman belajar yang lebih personal dan efisien.

2. DATA DAN METODOLOGI

Jenis penelitian yang digunakan dalam studi ini adalah penelitian kuantitatif, yang bertujuan untuk mengukur dan menganalisis data numerik secara sistematis untuk memperoleh kesimpulan yang objektif dan terukur. Pendekatan yang dipilih adalah uji t/Wilcoxon satu sampel, yaitu metode statistik yang digunakan untuk menguji apakah rata-rata suatu sampel berbeda secara signifikan dari nilai rata-rata populasi (μ_0) yang telah ditetapkan sebelumnya. Metode ini tepat digunakan ketika data yang dikumpulkan berasal dari satu kelompok

sampel dan ingin dibandingkan dengan nilai acuan tertentu. Pemilihan metode ini didasarkan pada sedikitnya jumlah sampel (<30) serta adanya kebutuhan untuk menguji hipotesis terkait perbedaan rata-rata dengan populasi tanpa memerlukan data dari kelompok pembandingan lain. Kajian uji t satu sampel ini dilakukan untuk menguji pertanyaan: apakah rata-rata kemampuan dan penggunaan teknologi kecerdasan buatan aplikasi AI-G oleh mahasiswa program studi Informatika di Universitas Bhakti Asih Tangerang berbeda secara signifikan dari nilai acuan. Studi ini menetapkan nilai 3 sebagai nilai acuan uji t satu sampel. Nilai dipilih karena sesuai dengan asumsi nilai rata-rata kemampuan sampel dengan nilai hipotesis populasi (μ_0) pada penelitian ini.

Data

Kajian ini melibatkan mahasiswa dari Program Studi Informatika di Universitas Bhakti Asih Tangerang sebagai populasi. Sebanyak 11 mahasiswa dipilih sebagai sampel menggunakan metode nonprobabilitas (*purposive sampling*), yaitu mahasiswa yang telah memanfaatkan aplikasi AI-G dalam kegiatan akademik mereka. Walaupun ukuran sampel ini tergolong kecil, penelitian metodologis menunjukkan bahwa uji t satu sampel masih dapat digunakan pada ukuran sampel yang sangat kecil, asalkan efek yang diharapkan cukup signifikan dan asumsi distribusi normal relatif terpenuhi (de Winter, 2013). Data dari 11 mahasiswa berhasil didapatkan (**Tabel 1**). Nama responden disamarkan menjadi huruf. Pertanyaan 1, 2, dan 3 diberi kode G01, G02, dan G03. Periode pengambilan data adalah minggu ke-2–3 bulan Februari 2026. Permasalahan yang dialami adalah lambatnya respons dari mahasiswa.

Tabel 1. Data survei yang diajukan kepada responden mahasiswa Prodi Informatika.

No	Responden	AI-G		
		G01	G02	G03
1	A	1	3	3
2	B	3	3	3
3	C	1	3	3
4	D	3	3	3
5	E	2	3	2
6	F	3	1	4
7	G	2	2	2
8	H	5	4	3
9	I	1	3	3
10	J	1	3	3
11	K	3	4	3

Metodologi

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan uji t satu sampel/Wilcoxon, jika data tidak normal, untuk menguji hipotesis bahwa rata-rata kemampuan dan penggunaan teknologi kecerdasan buatan, aplikasi AI-G mahasiswa program studi Informatika di Universitas Bhakti Asih Tangerang berbeda secara signifikan dari nilai acuan yang telah ditetapkan. Uji t satu sampel dipilih karena sesuai untuk membandingkan nilai rata-rata sampel dengan sebuah nilai hipotesis populasi (μ_0).

Populasi penelitian adalah seluruh mahasiswa aktif Program Studi Informatika di Universitas Bhakti Asih Tangerang. Sampel dipilih sebanyak 11 mahasiswa melalui metode nonprobabilitas (*purposive sampling*), yakni mahasiswa yang telah menggunakan aplikasi AI dalam pembelajaran akademiknya. Meskipun ukuran sampel relatif kecil, studi metodologis menyatakan bahwa Uji-T Satu Sampel dapat tetap diterapkan pada ukuran sangat kecil, namun dengan catatan efek yang diharapkan cukup besar dan asumsi distribusi normal relatif terpenuhi (de Winter, 2013).

Data dari 11 mahasiswa Program Studi Informatika di Universitas Bhakti Asih Tangerang telah berhasil dikumpulkan. Mereka adalah mahasiswa semester 5 yang dianggap sudah memiliki pengetahuan dasar informatika yang memadai untuk menjawab pertanyaan survei dengan baik. Pengumpulan data dilakukan secara daring melalui *Google Docs*, di mana responden mengisi kolom dalam file yang disediakan. Pertanyaan utamanya adalah: 1) Apakah Anda sering menggunakan aplikasi AI-G untuk membantu kegiatan akademik, terutama pembuatan salindia? 2) Apakah aplikasi AI-G bermanfaat dalam kegiatan akademik yang Anda lakukan? 3) Apakah Anda merasa nyaman (mudah) menggunakan aplikasi AI-G sebagai alat bantu pembelajaran? Pengukuran dilakukan dengan skala Likert 5 poin (1 = sangat rendah hingga 5 = sangat tinggi; **Tabel 2**). Sebelum analisis, dilakukan uji normalitas sampel melalui uji Shapiro-Wilk. Data yang terkumpul kemudian dianalisis menggunakan software JAMOWI untuk menghitung rata-rata ($\bar{x}$), simpangan baku (s), dan standar error ($s/\sqrt{n}$).

Hipotesis nol (H_0): rata-rata kemampuan dan penggunaan aplikasi AI-G mahasiswa = nilai acuan (μ_0). Hipotesis alternatif (H_1): rata-rata berbeda dari μ_0 (dua arah). Uji satu sampel t statistik dihitung sebagai:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

dengan derajat kebebasan (df) = $n-1=10$

Tabel 2. Skala Likert ini digunakan dan diajukan kepada mahasiswa sebagai respons terhadap setiap jawaban di dalam kuesioner.

No.	Kode	Keterangan	Likert
1.	SS	Sangat Setuju	5
2.	S	Setuju	4
3.	N	Netral	3
4.	TS	Tidak Setuju	2
5.	STS	Sangat Tidak Setuju	1

Hipotesis nol (H_0): rata-rata kemampuan dan penggunaan AI mahasiswa = nilai acuan (μ_0).
Hipotesis alternatif (H_1): rata-rata berbeda dari μ_0 (dua arah). Uji satu sampel t-statistik dihitung sebagai:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

dengan derajat kebebasan (df) = $n-1=10$

Kode Etik

Peserta diberikan penjelasan tentang tujuan penelitian dan dijamin kerahasiaan identitas sebagai bentuk perhatian terhadap kode etik penelitian. Kajian ini menerapkan perlindungan data sesuai standar akademik. Peserta mahasiswa diyakinkan bahwa apa pun jawaban mereka tidak berhubungan dengan kinerja akademiknya. Ini adalah sebuah penelitian. Data yang diinput dalam sistem dan *file*/berkas tidak menyebutkan nama siswa secara langsung dan hanya menggunakan kode huruf. Pengambilan data dilakukan pada minggu ke-2 dan ke-3 bulan Februari 2026. Kendala yang dihadapi umumnya adalah keterlambatan jawaban dari responden.

Penulis menyatakan tidak memiliki konflik kepentingan atau *conflict of interest* dalam kajian ini. Operasional studi ini berasal dari pendanaan internal Universitas Bhakti Asih Tangerang yang dituangkan dalam SK penugasan penelitian dengan nomor 125-J/SK/UNIBANG.REK/I/2026.

Perangkat

Pengolahan data dilakukan dengan perangkat JAMOWI 2.6. Jamovi adalah perangkat lunak analisis statistik *open-source* yang dirancang sebagai alternatif gratis, *open-source* dan modern untuk program komersial seperti SPSS atau SAS. Secara khusus, Jamovi versi 2.6 diluncurkan pada awal tahun 2025 dengan berbagai pembaruan pada stabilitas dan fungsi modular. Jamovi diinisiasi oleh tiga pengembang inti: Jonathon Love, Damian Dropmann, dan Ravi Selker. Perangkat lunak ini pertama kali diperkenalkan ke publik dan komunitas akademis pada awal tahun 2017 di University of Newcastle, Australia. Perangkat lunak ini telah diadopsi secara resmi oleh lebih dari 300 universitas di seluruh dunia sebagai kurikulum standar untuk mengajarkan statistik (Langkamp, 2024).

3. HASIL PENELITIAN

Statistik Deskriptif

Pengujian statistik deskriptif dilakukan untuk pertanyaan 1 (G01), 2 (G02), dan 3 (G03) secara terpisah (**Tabel 3**). Hal ini dilakukan untuk mengantisipasi perbedaan sifat normalitas data. Berdasarkan hasil analisis statistik deskriptif terhadap tiga kelompok variabel dengan jumlah sampel masing-masing sebanyak sebelas responden dan tanpa adanya data yang hilang, terlihat adanya karakteristik distribusi data yang cukup beragam. Dari sisi ukuran tendensi sentral, pertanyaan G02 dan G03 menunjukkan nilai rata-rata sedang, yakni masing-masing sebesar 2,91 dengan nilai median 3,00. Angka ini mengindikasikan kecenderungan skor yang lebih moderat dibandingkan dengan kelompok G01 yang memiliki rata-rata lebih rendah, yaitu 2,27 dengan median 2,00.

Tabel 3. Statistik deskriptif dari data penelitian.

	N	Mean	Median	SD	SE
G01	11	2.27	2	1.272	0.384
G02	11	2.91	3	0.831	0.251
G03	11	2.91	3	0.539	0.163

Uji Normalitas

Secara metodologis, mengingat ukuran sampel yang tergolong kecil, yaitu di bawah tiga puluh, asumsi normalitas data menjadi prasyarat mutlak yang harus diuji terlebih dahulu sebelum melangkah ke tahap analisis inferensial. Uji normalitas pada kajian ini mengacu pada tes Shapiro-Wilk (**Tabel 4**) untuk mengetahui apakah distribusi data mengikuti kaidah normalitas atau sebaliknya (Shapiro & Wilk, 1965). Berdasarkan pengujian Shapiro-Wilk terhadap ketiga data, diperoleh informasi mengenai sebaran data dari masing-masing variabel dengan ukuran sampel yang konsisten sebanyak sebelas responden per kelompok tanpa adanya data yang hilang. Pada kelompok G01, nilai statistik uji Shapiro-Wilk tercatat sebesar 0,851 dengan nilai p-value sebesar 0,044. Selanjutnya, pada kelompok G02, diperoleh nilai statistik uji sebesar 0,795 dengan p-value yang lebih kecil, yaitu 0,008. Sementara itu, kelompok G03 menunjukkan nilai statistik uji sebesar 0,724 dengan p-value yang sangat kecil atau kurang dari 0,001. Mengingat seluruh nilai p-value pada ketiga kelompok tersebut berada di bawah tingkat signifikansi standar sebesar 0,05, dapat disimpulkan secara statistik bahwa data dari kelompok G01, G02, dan G03 tidak berdistribusi normal. Oleh karena itu, dalam tahapan analisis inferensial atau uji hipotesis lanjutan, penggunaan metode statistik parametrik uji Student's t dihindari, sehingga pada kajian ini metode yang digunakan adalah statistika nonparametrik Wilcoxon (Arnold, 1965; Minitab Blog Editor, 2016), agar hasil analisis tetap valid dan dapat dipertanggungjawabkan secara ilmiah.

Tabel 4. Hasil uji normalitas data penelitian.

		W statistic	p
G01	Shapiro-Wilk	0.851	0.044
G02	Shapiro-Wilk	0.795	0.008
G03	Shapiro-Wilk	0.724	<.001

Hasil Uji Wilcoxon Signed Rank

Berdasarkan hasil uji satu sampel nonparametrik menggunakan metode *Wilcoxon Signed-Rank* dengan hipotesis alternatif bahwa nilai rata-rata populasi tidak sama dengan tiga, diperoleh temuan dari masing-masing kelompok dengan ukuran sampel sebanyak sebelas responden (**Tabel 5**). Pada kelompok G01, nilai statistik Wilcoxon W tercatat sebesar 5,000 dengan nilai p-value sebesar 0,135. Selanjutnya, pada kelompok G02, diperoleh nilai statistik Wilcoxon W sebesar 4,000 dengan p-value sebesar 0,850. Sementara itu, kelompok

G03 menunjukkan nilai statistik Wilcoxon W sebesar 2,000 dengan p-value sebesar 0,773. Karena seluruh nilai p-value pada ketiga kelompok tersebut berada di atas tingkat signifikansi standar sebesar 0,05, dapat disimpulkan bahwa tidak terdapat perbedaan yang signifikan secara statistik antara nilai median atau rata-rata dari kelompok G01, G02, dan G03 dengan nilai acuan tiga.

Tabel 5. Hasil uji Wilcoxon satu sampel pada penelitian ini.

		Statistic	p	Mean difference	Effect Size
G01	Wilcoxon W	5.00	0.135	-1.500	-0.643
G02	Wilcoxon W	4.00	0.850	-0.353	-0.200
G03	Wilcoxon W	2.00	0.773	-0.172	-0.333

4. PEMBAHASAN

Kuesioner 1

Hasil uji Wilcoxon Signed-Rank Test pada pertanyaan G01 menunjukkan bahwa secara statistik tidak terdapat perbedaan yang signifikan jika dibandingkan dengan nilai acuan 3, dilihat dari p-value sebesar 0,135 yang lebih besar dari tingkat signifikansi 0,05. Hal ini mengindikasikan bahwa data yang dianalisis tidak cukup kuat untuk menolak hipotesis nol, sehingga secara formal tidak dapat disimpulkan adanya perubahan atau perbedaan yang bermakna pada kelompok tersebut. Namun, penting untuk dicatat bahwa uji ini sangat dipengaruhi oleh ukuran sampel yang kecil, sehingga kekuatan statistik (statistical power) uji menjadi terbatas dan berpotensi menyebabkan hasil yang kurang sensitif terhadap perbedaan yang sebenarnya ada.

Meski demikian, aspek substantif dari analisis ini menunjukkan adanya kecenderungan empiris yang menarik, yaitu *mean difference* sebesar -1,500 yang berarti skor kelompok G01 cenderung lebih rendah dibandingkan dengan nilai acuan. Ukuran efek (*effect size*) yang diukur dengan *rank biserial correlation* sebesar -0,643 menunjukkan kekuatan hubungan yang besar, yang mengimplikasikan bahwa meskipun hasil uji statistik tidak signifikan secara formal, secara praktis terdapat indikasi perbedaan yang relevan.

Kuesioner 2

Pada pertanyaan G02, hasil uji Wilcoxon dengan nilai statistik W sebesar 4,000 dan p-value sebesar 0,850 menunjukkan bahwa tidak terdapat perbedaan signifikan antara data kelompok ini dan nilai acuan 3. P-value yang jauh lebih besar dari batas α 0,05 mengindikasikan bahwa hipotesis nol tidak dapat ditolak, sehingga dapat disimpulkan bahwa nilai median kelompok G02 tidak berbeda secara bermakna dari nilai acuan tersebut. Hal ini diperkuat oleh *mean difference* yang relatif kecil sebesar -0,353, yang mengindikasikan perbedaan rata-rata yang minimal dan tidak signifikan secara statistik.

Ukuran efek yang diukur menggunakan *rank biserial correlation* sebesar -0,200 memperlihatkan bahwa perbedaan antara kelompok G02 dan nilai acuan memiliki magnitudo yang sedang, sehingga pengaruh praktisnya dalam konteks evaluasi sistem ini

dapat dianggap menengah/cukup besar. Dengan demikian, hasil ini menegaskan bahwa performa atau karakteristik kelompok G02 cukup berbeda dari standar atau nilai acuan yang telah ditetapkan, tanpa adanya indikasi perubahan atau perbedaan yang signifikan.

Kuesioner 3

Analisis pada kelompok G03 menunjukkan bahwa nilai statistik Wilcoxon sebesar 2,000 dengan p-value 0,773 menegaskan tidak adanya perbedaan signifikan secara statistik antara data kelompok tersebut dan nilai acuan 3. Hal ini konsisten dengan hasil pada kelompok lain, yang menunjukkan bahwa distribusi data dalam kelompok G03 tidak berbeda secara bermakna dari nilai referensi yang ditetapkan. Meskipun *mean difference* yang tercatat paling kecil yaitu -0,172, serta ukuran efek *rank biserial correlation* sebesar -0,333 menunjukkan deviasi yang relatif sedang, hal ini tetap memberikan gambaran adanya perbedaan menengah yang cukup signifikan secara statistik dalam konteks pengujian hipotesis.

Dari perspektif metodologi, meskipun ketiga kelompok tidak menunjukkan perbedaan signifikan terhadap nilai acuan, perhatian khusus harus diberikan pada ukuran efek yang cukup besar pada kelompok G01, yang dapat mengindikasikan potensi pengaruh yang belum terdeteksi secara statistik akibat keterbatasan ukuran sampel. Oleh karena itu, rekomendasi utama untuk penelitian lanjutan di bidang informatika adalah peningkatan jumlah sampel guna memperkuat validitas dan sensitivitas uji statistik yang digunakan. Dengan demikian, penambahan sampel diharapkan dapat meningkatkan stabilitas hasil dan memberikan gambaran yang lebih akurat mengenai perbedaan yang mungkin ada, sekaligus mendukung pengambilan kesimpulan yang lebih kuat dan terpercaya dalam konteks penelitian serupa.

Diskusi

Hasil penelitian menunjukkan bahwa penggunaan aplikasi AI-G di lingkungan responden masih kurang. Hal ini kemungkinan menunjukkan bahwa mahasiswa tidak bergantung pada aplikasi AI untuk pembuatan salindia, tetapi masih membuatnya secara manual dengan aplikasi konvensional seperti Ms Office. Dilihat dari sudut pandang yang lain, dapat dikatakan bahwa mahasiswa memiliki literasi yang kurang dalam pemanfaatan AI untuk pembuatan salindia. Jika ini yang terjadi, maka pendidik wajib memperkenalkan aplikasi-aplikasi AI yang ada di dunia akademik untuk meningkatkan literasi digital mahasiswa. Zaman disrupsi AI yang sedang berlangsung saat ini tidak bisa dihindari; hanya mereka yang sanggup beradaptasi dengan AI yang akan mencapai puncak kesuksesan. Seperti kata Jensen Huang, CEO NVIDIA, ketika berkunjung ke Indonesia pada November 2024, AI tidak akan menggantikan manusia, tetapi manusia yang tidak menggunakan AI akan digantikan oleh manusia yang menggunakan AI (Taufiqurrahman, 2026).

5. KESIMPULAN

Sesuai dengan pembahasan di bagian pendahuluan. Penelitian ini bertujuan untuk membandingkan tingkat pemakaian aplikasi AI-G oleh mahasiswa Program Studi Informatika di Universitas Bhakti Asih Tangerang (UNIBANG) dengan nilai standar acuan (3). Hal ini direpresentasikan dalam (1) mengukur rerata kemampuan mahasiswa dalam memanfaatkan Aplikasi Gamma App untuk keperluan akademik yang dipresentasikan melalui tiga pertanyaan atau kuesioner, (2) menguji apakah nilai tersebut secara statistik

berbeda dengan nilai acuan yang telah ditetapkan, Pada penelitian ini digunakan angka 3.

Berdasarkan data statistik deskriptif, dapat diketahui bahwa nilai rerata mahasiswa untuk penggunaan Gamma App yang direpresentasikan oleh tiga pertanyaan, menunjukkan nilai 2,27 untuk pertanyaan G01 dan 2,91 untuk pertanyaan G02 dan G03. Artinya, hampir tidak ada perbedaan dengan nilai acuan 3, bahkan rerata kemampuan mahasiswa cenderung berada di bawah nilai acuan.

Data kuesioner G01, G02, dan G03 memiliki pola distribusi yang tidak normal. Hal ini dapat dilihat pada nilai uji Shapiro-Wilk yang menunjukkan nilai p-value pada G01 = 0,044. Sedangkan kuesioner G02 dan G03 menunjukkan nilai 0,008 dan 0,001. Semua ini menunjukkan ketidaknormalan. Nilai *p* yang sangat kecil, yaitu lebih kecil dari 0,05. Ketidaknormalan ini menunjukkan bahwa data hanya dapat dianalisis dengan pengujian nonparametrik seperti Wilcoxon signed-rank satu sampel.

Untuk kelompok G01, statistik Wilcoxon W tercatat sebesar 5,000, dengan nilai *p* sebesar 0,135. Selanjutnya, untuk kelompok G02, statistik Wilcoxon W adalah 4,000, dengan nilai *p* sebesar 0,850. Sementara itu, kelompok G03 menunjukkan statistik Wilcoxon W sebesar 2,000, dengan nilai *p* sebesar 0,773. Karena nilai *p* untuk ketiga kelompok tersebut melampaui tingkat signifikansi standar 0,05, dapat disimpulkan bahwa tidak terdapat perbedaan yang signifikan secara statistik antara nilai median atau rata-rata kelompok G01, G02, dan G03 dengan nilai acuan tiga.

Kelemahan Penelitian dan Kajian Lanjutan

Penelitian ini menggunakan ukuran sampel sebanyak 11 per kelompok. Untuk memperoleh hasil yang lebih representatif dan meningkatkan kekuatan statistik, penambahan ukuran sampel sangat disarankan. Data menunjukkan perbedaan yang kecil namun konsisten, yang mengindikasikan tingkat reliabilitas yang memadai. Penelitian lebih lanjut dapat menggunakan analisis regresi atau multivariat untuk memperoleh wawasan lebih mendalam mengenai faktor-faktor yang memengaruhi hasil—seperti dampak variabel eksternal yang tidak terdeteksi oleh uji sampel tunggal.

Desain penelitian yang menggunakan kuesioner dapat diawali dengan uji reliabilitas. Uji ini menilai keandalan hasil yang diperoleh dari instrumen—khususnya, apakah hasil tersebut tetap konsisten ketika pengujian diulang dalam kondisi serupa. Skor reliabilitas yang tinggi menunjukkan bahwa instrumen pengukuran tersebut dapat diandalkan untuk menghasilkan hasil yang konsisten.

Salah satu contoh uji reliabilitas adalah *Split-Half Reliability*, yang melibatkan pembagian instrumen menjadi dua bagian (misalnya, membagi butir pertanyaan menjadi dua) dan menghitung korelasi antara hasil dari kedua bagian tersebut. Korelasi yang tinggi menunjukkan bahwa instrumen tersebut konsisten. Metode reliabilitas lainnya adalah Cronbach's Alpha (α), metode yang paling umum digunakan untuk mengukur reliabilitas instrumen yang terdiri dari beberapa butir soal atau pertanyaan. Nilai Alpha berkisar antara 0 hingga 1; nilai yang mendekati 1 menunjukkan reliabilitas yang lebih tinggi, dengan ambang batas minimum yang dapat diterima sebesar 0,7 (Setyaedhi, 2024).

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Bhakti Asih Tangerang yang telah memberikan dukungan pendanaan internal untuk menjalankan penelitian ini serta rekan sejawat yang telah membantu dalam penulisan manuskrip karya ilmiah.

PUSTAKA

- Abbasnasab Sardareh, S., Brown, G. T. L., & Denny, P. (2021). Comparing four contemporary statistical software tools for introductory data science and statistics in the social sciences. *Teaching Statistics*, 43(S1).
<https://doi.org/10.1111/test.12274>
- Arnold, H. J. (1965). Small Sample Power of the One Sample Wilcoxon Test for Non-Normal Shift Alternatives. *The Annals of Mathematical Statistics*, 36(6), 1767–1778.
<https://doi.org/10.1214/aoms/1177699805>
- Box, J. F. (1987). Guinness, Gosset, Fisher, and Small Samples. *Statistical Science*, 2(1).
<https://doi.org/10.1214/ss/1177013437>
- Destiani, R. D., & Mufiidah, A. N. (2024). Era Baru Ekonomi Digital: Studi Komprehensif tentang Teknologi dan Pasar. *ADI Bisnis Digital Interdisiplin Jurnal*, 5(1), 47–50.
<https://doi.org/10.34306/abdi.v5i1.1095>
- Deus, D. M. D., Carvalho, C. P. D., Mazza, F. C., Siqueira, S. S. S. D., Miranda Junior, J. E., & Sampaio, N. A. D. S. (2024). Using Analysis of Variance and Tukey's Test to Compare Four Brands of Industrial Paint. *Revista de Gestão Social e Ambiental*, 18(2), e07582.
<https://doi.org/10.24857/rgsa.v18n2-161>
- Dikkatwar, R., Mohite, M., Sharma, S., Selvakumar, P., Jayadeva, S. M., & T C, M. (2025). Artificial Intelligence and Adaptive Learning Technologies (pp. 313–338). Igi Global Scientific.
<https://doi.org/10.4018/979-8-3373-5157-5.ch011>
- Harris, T., & Hardin, J. W. (2013). Exact Wilcoxon Signed-Rank and Wilcoxon Mann–Whitney Ranksum Tests. *The Stata Journal: Promoting Communications on Statistics and Stata*, 13(2), 337–343.
<https://doi.org/10.1177/1536867x1301300208>
- Ichikawa, M., & Konishi, S. (1995). Application of the Bootstrap Methods in Factor Analysis. *Psychometrika*, 60(1), 77–93.
<https://doi.org/10.1007/bf02294430>
- Jankowski, K. R. B., Flannelly, K. J., & Flannelly, L. T. (2017). The t-test: An Influential Inferential Tool in Chaplaincy and Other Healthcare Research. *Journal of Health Care Chaplaincy*, 24(1), 30–39.
<https://doi.org/10.1080/08854726.2017.1335050>
- Langkamp, T. (2024). Farewell SPSS: Switching to free software in 2024. Center for Open Science.
<https://doi.org/10.31234/osf.io/huajb>
- Mcgee, M. (2018). Case for omitting tied observations in the two-sample t-test and the Wilcoxon-Mann-Whitney Test. *PloS One*, 13(7), e0200837.
<https://doi.org/10.1371/journal.pone.0200837>
- Minitab Blog Editor. (2016, August 22). *Data Not Normal? Try Letting It Be, with a*

- Nonparametric Hypothesis Test*. Minitab, LLC. Blog.Minitab.Com. <https://blog.minitab.com/en/blog/understanding-statistics/data-not-normal-try-letting-it-be-with-a-nonparametric-hypothesis-test>
- Mustafidah, H., Imantoyo, A., & Suwarsito, S. (2020). Pengembangan Aplikasi Uji-t Satu Sampel Berbasis Web. *JUITA: Jurnal Informatika*, 8(2), 245. <https://doi.org/10.30595/juita.v8i2.8786>
- Nahm, F. S. (2016). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*, 69(1), 8. <https://doi.org/10.4097/kjae.2016.69.1.8>
- Navarro, D., & Foxcroft, D. (2025). *Learning Statistics with jamovi*. Open Book. <https://doi.org/10.11647/obp.0333>
- Noordzij, M., Tripepi, G., Dekker, F. W., Zoccali, C., Tanck, M. W., & Jager, K. J. (2010). Sample size calculations: basic principles and common pitfalls. *Nephrology Dialysis Transplantation*, 25(5), 1388–1393. <https://doi.org/10.1093/ndt/gfp732>
- Panagiotou, N. (2025). AI Literacy and the Future of Education: A Framework for Ethical and Inclusive Learning Models. *Journal of Interdisciplinary Knowledge*, 8, e01635. <https://doi.org/10.37497/jik.v8iknowledge.1635>
- Rao, C. R. (1992). R. A. Fisher: The Founder of Modern Statistics. *Statistical Science*, 7(1). <https://doi.org/10.1214/ss/1177011442>
- Rodzi, Z. M., Rahman, A. A., Razali, I. N., Nazri, I. S. M., & Abd Gani, A. F. (2023). Unraveling the Drivers of Artificial Intelligence (AI) Adoption in Higher Education. 1–6. <https://doi.org/10.1109/incult59088.2023.10482651>
- Ross, A., & Willson, V. L. (2017). One-Sample T-Test (pp. 9–12). Sensepublishers. https://doi.org/10.1007/978-94-6351-086-8_2
- Şahin, M., & Aybek, E. (2020). Jamovi: An Easy to Use Statistical Software for the Social Scientists. *International Journal of Assessment Tools in Education*, 6(4), 670–692. <https://doi.org/10.21449/ijate.661803>
- Setyaedhi, H. S. (2024). Comparative Test of Cronbach's Alpha Reliability Coefficient, Kr-20, Kr-21, And Split-Half Method. *Journal of Education Research and Evaluation*, 8(1), 47–57. <https://doi.org/10.23887/jere.v8i1.68164>
- Taufiqurrahman, A. N. (2026). Literasi Digital Untuk Penelitian Akademik. In *Literasi Digital Dalam Dunia Akademik* (1st ed., p. 156). PT. Literatus Digitus Indonesia. <https://www.lidigin.com/product/literasi-digital-dalam-dunia-akademik/>
- Zabell, S. L. (2008). On Student's 1908 Article "The Probable Error of a Mean." *Journal of the American Statistical Association*, 103(481), 1–7. <https://doi.org/10.1198/016214508000000030>
- Zhang, Z. (2013). Monte Carlo based statistical power analysis for mediation models: methods and software. *Behavior Research Methods*, 46(4), 1184–1198. <https://doi.org/10.3758/s13428-013-0424-0/>