

ANALISIS KESALAHAN SISWA DALAM MENYELESAIKAN SOAL ANALISIS KORELASI DI TINJAU DARI KESALAHAN KONSEP DAN PROSES

Ani Jamiatun Aisyah¹, Sutini², Agung Prasetyo³

^{1,2,3} Fakultas Tarbiyah dan Keguruan, Universitas Islam Negeri Sunan Ampel Surabaya
Email : anijamiatun381@email.com, sutinimiskun@uinsa.ac.id, agung.prasetyo@uinsa.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menganalisis kesalahan siswa dalam menyelesaikan soal analisis korelasi ditinjau dari kesalahan konsep dan proses. Latar belakang penelitian ini berangkat dari kenyataan bahwa banyak siswa masih kesulitan memahami materi statistika, khususnya analisis korelasi, sehingga rentan melakukan kesalahan baik pada penggunaan rumus, perhitungan, maupun interpretasi hasil. Penelitian ini menggunakan pendekatan deskriptif kuantitatif dengan subjek 36 siswa kelas XI SMA Negeri 3 Sidoarjo. Instrumen penelitian berupa tes uraian empat soal mengenai korelasi Pearson dan koefisien determinasi. Analisis kesalahan dilakukan berdasarkan kriteria Watson yang mencakup kesalahan data, prosedur, hingga kesimpulan. Hasil penelitian menunjukkan bahwa 50% siswa mengalami kesalahan dalam menghitung koefisien korelasi Pearson (r), 11% tidak menjawab, dan hanya 39% yang benar. Kesalahan umum meliputi kesalahan substitusi, perhitungan SS_{xy} , SS_{xx} , dan SS_{yy} , serta kesalahan dalam menginterpretasi nilai koefisien determinasi (r^2). Kesimpulan penelitian ini menegaskan bahwa kesalahan siswa mencakup aspek konsep dan proses, sehingga guru perlu memberikan pembelajaran berbasis pemahaman konsep, latihan bertahap, dan penggunaan data kontekstual.

Kata Kunci: *Analisis Kesalahan, Korelasi, Konsep, Proses*

ABSTRACT

This study aims to analyze students' errors in solving correlation analysis problems in terms of conceptual and procedural mistakes. The background arises from the fact that many students struggle to understand statistics, especially correlation analysis, which often leads to errors in applying formulas, performing calculations, and interpreting results. This research employed a quantitative descriptive approach involving 36 eleventh-grade students of SMA Negeri 3 Sidoarjo. The instrument was a written test consisting of four essay questions on Pearson correlation and coefficient of determination. Error analysis was based on Watson's criteria, including data usage, procedure, and conclusion errors. The findings revealed that 50% of students made mistakes in calculating the Pearson correlation coefficient (r), 11% did not answer, and only 39% answered correctly. Common mistakes included substitution errors, miscalculations of SS_{xy} , SS_{xx} , and SS_{yy} , and misinterpretation of the coefficient of determination (r^2). The study concludes that students' errors cover both conceptual and procedural aspects, suggesting that teachers should emphasize conceptual understanding, gradual practice, and contextual data in teaching correlation analysis.

Keywords: *Error analysis, Correlation, Concept, Process*

PENDAHULUAN

Matematika merupakan sebuah disiplin ilmu yang unik, dicirikan oleh sifatnya yang deduktif, sistematis, dan sangat terstruktur. Fondasi ilmu ini dibangun di atas kesepakatan logis berupa definisi, aksioma, serta teorema-teorema yang saling berkaitan erat dan membentuk sebuah sistem yang koheren. Dalam konteks pembelajaran, penguasaan matematika tidak hanya menekankan pada kemampuan berhitung atau keterampilan prosedural semata. Lebih dari itu, pembelajaran matematika menuntut siswa untuk mengembangkan pemahaman konsep yang

mendalam. Siswa diharapkan mampu mengaplikasikan proses penyelesaian yang benar secara logis dan sistematis dalam memecahkan berbagai permasalahan (Aswandari et al., 2025; Lestari et al., 2025). Oleh karena itu, keberhasilan siswa dalam mempelajari matematika tidak dapat diukur dari satu aspek saja, melainkan ditentukan oleh penguasaan yang seimbang antara pemahaman konseptual yang kokoh dan keterampilan prosedural yang akurat. Keduanya harus berjalan beriringan untuk mencapai kompetensi matematika yang utuh dan fungsional.

Namun pada kenyataannya, terdapat kesenjangan yang lebar antara idealisme tersebut dengan realitas di lapangan. Matematika secara konsisten masih menjadi salah satu mata pelajaran yang dianggap paling sulit oleh sebagian besar siswa. Karakteristik utama matematika yang bersifat abstrak, serta keterkaitannya yang sangat erat dengan logika deduktif, sering kali menimbulkan kesulitan signifikan bagi siswa dalam memahami materi. Konsep-konsep seperti variabel, fungsi, atau limit tidak memiliki wujud konkret sehingga sulit dibayangkan. Penelitian sebelumnya telah mengonfirmasi bahwa hambatan belajar (*learning barriers*) matematika ini dialami oleh siswa dari jenjang pendidikan dasar hingga menengah atas (Boyi & Rahayuningsih, 2025; Sholichah & Rahayuningsih, 2025). Faktor-faktor yang memengaruhinya pun beragam, mulai dari kurangnya kesiapan belajar siswa, rendahnya minat dan motivasi internal, serta keterbatasan pemahaman mereka terhadap penjelasan yang diberikan oleh guru (Jumrah et al., 2023). Akumulasi hambatan-hambatan tersebut berimplikasi langsung pada rendahnya pencapaian tujuan pembelajaran yang diharapkan.

Salah satu bidang dalam matematika yang terbukti cukup menantang bagi siswa adalah statistika. Sebagai cabang ilmu matematika, statistika memegang peranan yang sangat penting dalam kehidupan modern, terutama dalam hal pengumpulan data, pengolahan, analisis mendalam, serta proses penarikan kesimpulan yang valid dari sekumpulan data tersebut (Kusumaningpuri et al., 2022). Pada materi spesifik seperti analisis korelasi, siswa tidak hanya dituntut untuk menghafal rumus. Mereka harus mampu memahami esensi hubungan antara dua variabel, melakukan perhitungan teknis untuk menemukan nilai koefisien korelasi *Pearson*, dan yang tidak kalah penting, mampu menginterpretasikan makna dari hasil perhitungan tersebut. Kemampuan yang dituntut ini bersifat ganda, yakni tidak hanya bersifat teknis-prosedural, tetapi juga sangat konseptual. Konsekuensinya, jika pemahaman dasar siswa mengenai konsep hubungan variabel ini kurang kuat, maka kesalahan dalam prosedur maupun interpretasi akan sangat mudah terjadi.

Berbagai penelitian yang berfokus pada analisis kesalahan siswa menunjukkan bahwa kesalahan yang dilakukan siswa dalam menyelesaikan soal analisis korelasi umumnya dapat dikategorikan menjadi dua jenis utama, yaitu kesalahan konsep dan kesalahan proses. Kesalahan konsep terjadi ketika siswa tidak memahami makna, prinsip dasar, atau justifikasi teoretis di balik rumus yang mereka gunakan; mereka mungkin hafal rumus tetapi tidak tahu kapan atau mengapa menggunakanannya. Di sisi lain, kesalahan proses muncul ketika siswa keliru dalam menerapkan prosedur penyelesaian langkah demi langkah, meskipun mungkin secara samar mereka mengetahui konsep dasarnya. Untuk menganalisis jenis-jenis kesalahan ini secara lebih rinci, salah satu kerangka kerja yang banyak digunakan oleh peneliti adalah kriteria Watson (Damayanti, 2024; Winarto et al., 2024). Kriteria ini menyediakan klasifikasi komprehensif yang mencakup delapan jenis kesalahan, seperti kesalahan pengambilan data, kesalahan prosedur, kesalahan manipulasi aljabar, hingga kesalahan dalam penarikan kesimpulan akhir (Triyani & Yuhana, 2023; Lantang et al., 2021).

Seiring dengan pesatnya perkembangan era *digital*, pendekatan dalam menganalisis kesalahan belajar siswa pun mulai beralih. Analisis manual yang memakan waktu kini mulai dilengkapi dengan pendekatan berbasis teknologi yang lebih efisien dan mendalam. Salah satu inovasi yang potensial adalah pemanfaatan *data mining* dengan menggunakan algoritma

Apriori. Algoritma ini memiliki kemampuan untuk mengidentifikasi pola asosiasi atau keterkaitan antar item dalam sebuah *dataset* besar. Dalam konteks pendidikan, teknik ini dapat digunakan untuk menganalisis data jawaban siswa dan menemukan pola asosiasi antar jenis kesalahan. Misalnya, *data mining* dapat mengungkap bahwa siswa yang melakukan kesalahan dalam menentukan data (tipe A) cenderung juga melakukan kesalahan dalam penarikan kesimpulan (tipe B). Dengan teknik ini, guru dapat mengetahui pola kesalahan yang paling sering muncul secara bersamaan, sehingga strategi pembelajaran remedial dapat disesuaikan dengan kebutuhan spesifik siswa.

Berdasarkan hasil observasi awal yang dilakukan peneliti selama kegiatan asistensi mengajar di salah satu SMA Negeri di Sidoarjo, ditemukan bahwa permasalahan kesalahan siswa ini memang nyata terjadi. Para siswa teramat masih sering melakukan kesalahan dalam pengerjaan soal analisis korelasi. Kesalahan tersebut bervariasi, mulai dari tahap awal seperti kesalahan dalam menentukan data yang relevan (misalnya, salah mengidentifikasi variabel X dan Y), kesalahan dalam menerapkan prosedur perhitungan rumus *Pearson* yang panjang, maupun kesalahan fundamental dalam menarik kesimpulan yang tepat dari nilai koefisien yang diperoleh. Temuan empiris ini mempertegas adanya kesenjangan yang signifikan antara kemampuan ideal yang diharapkan oleh kurikulum dengan kenyataan kemampuan siswa di lapangan. Hal tersebut menjadi indikasi kuat perlunya analisis yang lebih mendalam terhadap jenis-jenis dan pola kesalahan yang dilakukan siswa agar guru dapat merancang intervensi pembelajaran yang tepat sasaran.

Dengan demikian, dilatarbelakangi oleh kesenjangan tersebut, penelitian ini memiliki dua tujuan utama. Pertama, penelitian ini bertujuan untuk menganalisis secara deskriptif jenis-jenis kesalahan siswa dalam menyelesaikan soal analisis korelasi, dengan fokus peninjauan berdasarkan kategori kesalahan konsep dan kesalahan proses. Kedua, sebagai nilai kebaruan (inovasi), penelitian ini juga mengkaji potensi penerapan pendekatan berbasis *data mining* untuk mengidentifikasi pola-pola kesalahan yang paling sering muncul secara bersamaan. Jika analisis kesalahan tradisional hanya mengidentifikasi *apa* kesalahannya, pendekatan ini diharapkan dapat mengungkap *bagaimana* satu kesalahan terkait dengan kesalahan lainnya. Melalui identifikasi pola ini, diharapkan penelitian ini dapat memberikan kontribusi praktis bagi guru dalam pengembangan strategi pembelajaran statistika yang lebih efektif, adaptif, dan sesuai dengan kebutuhan nyata siswa di kelas.

METODE PENELITIAN

Penelitian ini dirancang menggunakan pendekatan deskriptif kuantitatif. Pendekatan ini dipilih dengan tujuan utama untuk memetakan, mengidentifikasi, dan menganalisis secara sistematis kesalahan-kesalahan yang dilakukan siswa dalam menyelesaikan soal materi analisis korelasi. Fokus analisis ditinjau dari dua aspek utama, yaitu kesalahan konsep dan kesalahan proses. Penelitian ini tidak memberikan perlakuan atau intervensi, melainkan hanya mendeskripsikan frekuensi dan jenis kesalahan yang muncul secara alamiah setelah proses pembelajaran. Subjek penelitian yang dilibatkan adalah 36 siswa kelas XI SMA Negeri 3 Sidoarjo. Seluruh subjek ini merupakan satu rombongan belajar utuh yang dianggap telah menerima materi pembelajaran statistika, khususnya analisis korelasi, sehingga relevan untuk dianalisis kemampuannya. Pengumpulan data dilaksanakan pada rentang waktu yang telah ditentukan, di mana siswa diberikan instrumen tes dalam satu sesi pertemuan.

Instrumen utama yang digunakan untuk pengumpulan data adalah tes tertulis berbentuk uraian (esai). Tes ini terdiri atas empat soal yang dirancang secara khusus untuk mengukur kemampuan siswa dalam materi analisis korelasi. Soal-soal tersebut secara spesifik mencakup perhitungan untuk menemukan koefisien korelasi Pearson (r) dan perhitungan serta interpretasi

koefisien determinasi (r^2). Pemberian tes uraian ini bertujuan untuk melihat secara rinci proses penyelesaian dan alur berpikir siswa, tidak hanya jawaban akhir. Pelaksanaan tes dilakukan secara kolektif di dalam kelas, di mana seluruh 36 subjek penelitian mengerjakan soal yang sama dalam alokasi waktu 60 menit untuk memastikan kondisi pengerjaan yang seragam bagi semua siswa. Lembar jawaban siswa dari tes ini kemudian dikumpulkan sebagai data primer penelitian.

Data primer yang terkumpul berupa lembar jawaban siswa selanjutnya dianalisis menggunakan teknik analisis kesalahan (error analysis) secara kuantitatif deskriptif. Analisis ini difokuskan pada pengidentifikasi kesalahan konsep (pemahaman rumus dan interpretasi) dan kesalahan proses (prosedur dan perhitungan). Sebagai kerangka kerja, analisis kesalahan ini merujuk pada kriteria Watson (Triyani & Yuhana, 2023; Lantang et al., 2021). Kriteria ini digunakan untuk mengkategorikan jenis kesalahan yang spesifik, seperti kesalahan dalam penggunaan data (misalnya, salah substitusi nilai), kesalahan prosedur (misalnya, kesalahan dalam perhitungan SS_{xy} , SS_{xx} , dan SS_{yy}), data yang hilang (tidak menjawab), dan kesalahan dalam penarikan kesimpulan (misalnya, salah menginterpretasi nilai r^2). Frekuensi kemunculan setiap jenis kesalahan kemudian dihitung dan dipersentasekan untuk mendapatkan gambaran yang jelas mengenai kesalahan yang paling dominan dilakukan oleh siswa.

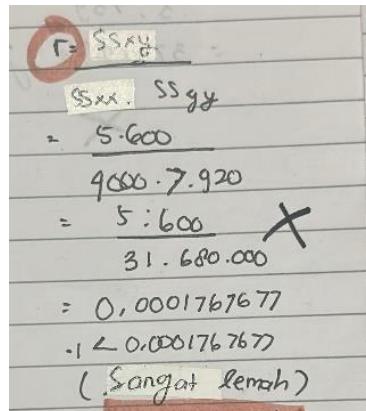
HASIL DAN PEMBAHASAN

Hasil

1. Kesalahan dalam Menentukan Nilai r (Koefisien Korelasi)

Berdasarkan hasil analisis mendalam terhadap lembar jawaban siswa, dapat diketahui bahwa pemahaman konsep terkait uji korelasi *Pearson* masih sangat terbatas dan muncul sejumlah miskonsepsi yang signifikan. Secara kuantitatif, dari 36 siswa yang berpartisipasi, hanya 14 siswa (sekitar 38,9%) yang berhasil menentukan nilai r dengan benar, sementara sebagian besar yakni 18 siswa (tepat 50%) melakukan kesalahan fatal dalam proses kalkulasi, dan 4 siswa lainnya (11,1%) memilih untuk tidak menjawab atau mengosongkan jawaban. Berdasarkan investigasi kualitatif atas jenis kesalahan yang dilakukan oleh 18 siswa tersebut, terdapat tiga akar permasalahan utama. Pertama, banyak siswa melakukan kesalahan substitusi nilai ke dalam rumus, sehingga terlihat adanya kebingungan dalam mengidentifikasi variabel seperti Σx , Σy , maupun n . Kedua, terdapat kesalahan perhitungan di langkah-langkah yang kompleks, yang menandakan kelemahan dalam akurasi aritmetika dasar mereka. Ketiga, dan yang paling mendasar, adalah ketidaktahuan siswa terhadap bentuk rumus korelasi *Pearson* itu sendiri, yang menunjukkan kegagalan dalam penguasaan konsep fundamental dan bukan sekadar kesalahan prosedural.

Contoh Kesalahan Siswa:


$$\begin{aligned} r &= \frac{SS_{xy}}{\sqrt{SS_{xx} \cdot SS_{yy}}} \\ SS_{xx} &= 5.600 \\ SS_{yy} &= 4000 \cdot 7.920 = 31.680.000 \\ r &= \frac{5.600}{31.680.000} = 0.0001767677 \\ r &= 0.0001767677 \\ &\text{(Sangat lemah)} \end{aligned}$$

Gambar 1. Kesalahan Substitusi

Gambar 1 menampilkan proses perhitungan r , yang merupakan koefisien korelasi antara dua variabel dengan menggunakan rumus rasio SS_{xy} dibagi akar dari hasil perkalian SS_{xx} dan SS_{yy} . Dalam tabel ini, SS_{xy} bernilai 5600, sementara perkalian SS_{xx} dan SS_{yy} diperoleh dari hasil 4000 dikalikan 7920 , yang sebenarnya memberikan hasil $31.680.000$, tetapi pada langkah berikutnya terdapat tanda silang yang menunjukkan adanya kesalahan pada tahap substitusi atau perhitungannya. Hasil akhir yang diperoleh adalah $0,0001767677$, dan angka tersebut menunjukkan bahwa nilai r sangat kecil atau mendekati nol, yang berarti hubungan antara dua variabel sangat lemah. Penulisan "sangat lemah" pada bagian akhir memperjelas interpretasi nilai koefisien korelasi yang telah dihitung. Kesalahan substitusi ini sangat penting untuk diidentifikasi agar analisis data menjadi akurat dan tidak menimbulkan kesalahan interpretasi pada hasil keseluruhan penelitian, sebagaimana yang diperlihatkan dalam

$$\begin{aligned}
 r &= \frac{SS_{xy}}{\sqrt{SS_{xx} \cdot SS_{yy}}} \\
 &= \frac{5.600}{\sqrt{4.000 \times 7.920}} \\
 &= \frac{5.600}{\cancel{100.909,78}} \quad \times \\
 &= \frac{5.600}{0,0111797695362} \\
 &= \text{sangat lemah}
 \end{aligned}$$

Gambar 2. Kesalahan hitung

Gambar 2 mengilustrasikan proses perhitungan koefisien korelasi dengan rumus $r = \frac{SS_{xy}}{\sqrt{SS_{xx} \cdot SS_{yy}}}$, namun ditemukan kekeliruan saat melakukan penghitungan akar dari hasil perkalian SS_{xx} dan SS_{yy} . Penulis menghitung akar dari 4000×7920 namun hasil yang didapat tertulis $100.909,78$, menunjukkan penyimpangan signifikan dari hasil seharusnya. Selanjutnya, hasil pembagian 5600 dengan 100.909,78 menghasilkan nilai $0,0111797695362$ yang tetap memberikan interpretasi hubungan "sangat lemah" antar variabel. Proses ini menjelaskan betapa krusialnya ketelitian dalam langkah penghitungan akar dan penggunaan rumus statistik secara akurat. Gambar 2 sangat berguna untuk menunjukkan dampak langsung dari kekeliruan dalam langkah perhitungan, karena dapat mempengaruhi interpretasi kekuatan hubungan data yang sedang diteliti, serta memberikan pembelajaran betapa pentingnya validasi hasil akhir dengan memeriksa kembali proses kalkulasi.

$$\begin{aligned}
 r &= \frac{5(11200) - 600(1130)}{\sqrt{5(76000) - 600(5(263300) - (1130)^2)}} \\
 r &= \frac{706000 - 678000}{\sqrt{38000 - 36000}} \quad (316500 - 276900) \\
 r &= \frac{28000}{\sqrt{2000 \times 39000}} = \sqrt{7920000} = 28132,61 \\
 r &= 0,995 \quad (\text{Positif Sangat Sama})
 \end{aligned}$$

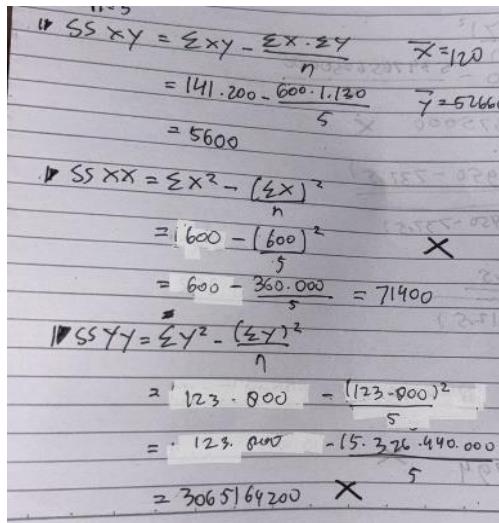
Gambar 3. Ketidaktahuan Rumus

Gambar 3 memperlihatkan berbagai upaya perhitungan koefisien korelasi dengan pendekatan rumus statistik yang berbeda, namun keseluruhannya kurang tepat. Penulis mencoba mengaplikasikan rumus dengan beberapa variasi, seperti menggunakan angka dan operasi matematika yang keliru pada tahap pengakaran dan penjumlahan, sehingga hasil akhirnya tidak konsisten dan kurang dapat diandalkan. Salah satu hasil yang dicatat adalah $r = 0,0995$ yang diberikan interpretasi sebagai hubungan positif namun sangat lemah. Gambar ini sangat penting untuk menunjukkan bagaimana ketidaktahan atau ketidakpahaman atas rumus atau tahapan penghitungan secara benar bisa menimbulkan kesalahan mendasar dalam interpretasi hasil analisis data. Dalam penelitian kuantitatif, penguasaan rumus serta pemahaman langkah penghitungan yang benar sangat krusial untuk mendapatkan hasil analisis yang valid, dan Gambar 3 menegaskan kebutuhan untuk pendidikan statistik yang mendalam dalam aplikasi nyata.

2. Kesalahan dalam Menghitung Nilai SSxy, SSxx, dan SSyy

Kesalahan sistematis kedua yang ditemukan dalam proses perhitungan korelasi adalah pada langkah krusial menghitung nilai *Sum of Squares* (SS), khususnya pada komponen SSxy, SSxx, dan SSyy. Kesalahan ini banyak terjadi karena siswa mengalami miskonsepsi dalam memahami sekaligus menerapkan rumus-rumus komputasi tersebut. Akarnya bersifat konseptual siswa sering keliru memahami simbol dalam rumus seperti gagal membedakan antara Σx^2 (jumlah dari data yang dikuadratkan) dan $(\Sigma x)^2$ (kuadrat dari total jumlah data), sehingga menyebabkan kesalahan saat menjumlahkan kuadrat data. Selain itu, ketidaktelitian dalam menyusun tabel bantu juga berkontribusi pada munculnya kesalahan baru; tabel yang seharusnya membantu justru menjadi sumber masalah karena pengisian data tidak dilakukan secara cermat. Kesalahan rangkaian ini sangat berdampak pada hasil akhir perhitungan r , dan secara berantai berpengaruh terhadap nilai r^2 (*koefisien determinasi*), bahkan menyebabkan interpretasi korelasi menjadi tidak valid secara statistik.

Contoh kesalahan siswa:



Handwritten calculations for SSxy, SSxx, and SSyy:

$$\begin{aligned} \text{1. } SS_{xy} &= \sum xy - \frac{\sum x \cdot \sum y}{n} \\ &= 141 \cdot 200 - \frac{600 \cdot 1.120}{5} \\ &= 5600 \end{aligned}$$

$$\begin{aligned} \text{2. } SS_{xx} &= \sum x^2 - \frac{(\sum x)^2}{n} \\ &= 1600 - \frac{(600)^2}{5} \\ &= 600 - \frac{360.000}{5} \\ &= 71400 \end{aligned}$$

$$\begin{aligned} \text{3. } SS_{yy} &= \sum y^2 - \frac{(\sum y)^2}{n} \\ &= 123 \cdot 800 - \frac{(123-800)^2}{5} \\ &= 123 \cdot 800 - 15.376.440.000 \\ &= 3065164200 \end{aligned}$$

Gambar 4. Kesalahan mencari nilai SSxx dan SSyy

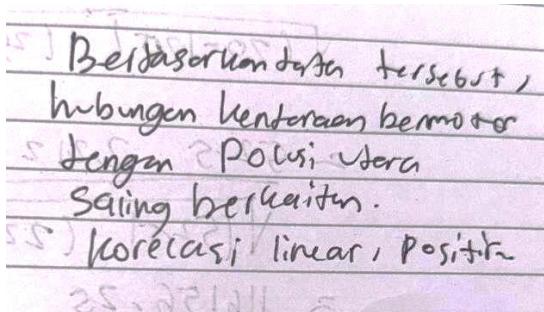
Gambar 4 menjelaskan rangkaian perhitungan komponen statistik yang penting dalam analisis regresi, yakni SSxy, SSxx, dan SSyy. Rumus yang digunakan sudah tepat, seperti pada SSxy yaitu dengan mengurangkan hasil perkalian jumlah xy dan hasil kali jumlah x dan jumlah y dibagi n, namun dalam langkah-langkah yang dilakukan untuk SSxx dan SSyy terdapat kesalahan terutama dalam proses mengkuadratkan jumlah total x dan y sebelum dibagi dengan n. Untuk SSxx, kesalahan terletak pada pengurangan $600 - 600^2/5$ yang ternyata menghasilkan nilai negatif tidak logis dalam statistik, dan untuk SSyy rumus yang dipakai juga menimbulkan Copyright (c) 2025 SCIENCE : Jurnal Inovasi Pendidikan Matematika dan IPA

hasil yang sangat besar dan tidak sesuai harapan. Perhitungan pada Gambar 4 ini menampilkan pentingnya pemahaman alur operasi matematika, khususnya proses pengkuadratan dan pembagian dalam statistik, agar tidak terjadi kesalahan signifikan yang dapat berakibat fatal pada hasil akhir analisis data dalam penelitian.

3. Kesalahan dalam Menentukan Interpretasi Berdasarkan Nilai r^2 (Koefisien Determinasi)

Kesalahan ketiga yang banyak ditemukan adalah pada tahap interpretasi data, khususnya dalam memahami hubungan antara nilai koefisien determinasi r^2 dengan nilai korelasi r . Banyak siswa tidak menyadari bahwa r^2 sepenuhnya bergantung pada nilai r yang telah dihitung sebelumnya, sehingga jika terjadi kesalahan pada perhitungan r , maka nilai r^2 pun akan otomatis salah. Akibatnya, interpretasi terhadap kekuatan hubungan antara variabel menjadi tidak valid dan dapat menyesatkan kesimpulan penelitian. Selain itu, beberapa siswa memberikan interpretasi hasil tanpa menyebutkan persentase kontribusi variabel bebas terhadap variabel terikat, padahal informasi tersebut penting untuk menilai seberapa besar pengaruh satu variabel terhadap variabel lainnya. Kesalahan lain yang sering terjadi adalah kekeliruan dalam menjelaskan makna hubungan antara variabel X dan Y, seperti menyatakan adanya hubungan sebab-akibat padahal analisis korelasi hanya menunjukkan hubungan linier tanpa menegaskan arah kausalitasnya.

Contoh kesalahan siswa:



Gambar 5. Interpretasi tanpa persentase

Gambar 5 menampilkan interpretasi akhir dari hasil analisis hubungan antara jumlah kendaraan bermotor dengan tingkat polusi udara tanpa menyebutkan persentase angka secara spesifik. Teks pada gambar menyatakan bahwa berdasarkan data yang telah dianalisis, kedua variabel tersebut saling berkaitan serta hubungan ini termasuk dalam korelasi linear yang bersifat positif. Artinya, apabila terjadi peningkatan pada jumlah kendaraan bermotor, maka tingkat polusi udara juga cenderung meningkat. Interpretasi ini dapat memberikan gambaran umum tanpa mendetailkan kekuatan dan besaran hubungan secara numerik, namun tetap menegaskan adanya keterkaitan sebab-akibat di antara kedua variabel tersebut secara umum.

Pembahasan

Analisis hasil tes terhadap 36 siswa SMA Negeri 3 Sidoarjo menyajikan temuan yang signifikan mengenai kesulitan belajar statistika, khususnya pada materi uji korelasi *Pearson*. Data kuantitatif menunjukkan gambaran yang mengkhawatirkan: hanya 38,9% (14 siswa) yang mampu menyelesaikan soal dengan benar, sementara mayoritas, tepatnya 50% (18 siswa), melakukan kesalahan dalam penggerjaannya, dan 11,1% sisanya memilih untuk tidak menjawab. Tingginya angka kesalahan ini mengindikasikan bahwa masalah yang dihadapi siswa bukanlah sekadar kekeliruan acak, melainkan bersifat sistematis dan mendasar. Temuan ini menegaskan bahwa pemahaman konseptual dan prosedural terkait analisis korelasi belum dikuasai oleh sebagian besar peserta didik, yang menandakan adanya tantangan serius dalam pembelajaran statistika di tingkat sekolah menengah.

Investigasi kualitatif terhadap jenis kesalahan dalam penentuan nilai r (koefisien korelasi) memperlihatkan masalah yang berlapis. Kesalahan yang teridentifikasi terbagi menjadi tiga kategori utama: kesalahan substitusi nilai ke dalam rumus (Gambar 1), kesalahan akurasi dalam perhitungan aritmetika (Gambar 2), dan yang paling fundamental adalah ketidaktahuan terhadap rumus korelasi *Pearson* itu sendiri (Gambar 3). Pola kesalahan ini mengimplikasikan bahwa kesulitan siswa tidak hanya terletak pada ketidakmampuan prosedural atau ketidaktelitian berhitung. Lebih dari itu, banyak siswa yang tampaknya tidak memiliki pemahaman konseptual dasar tentang apa yang mereka hitung. Ketidaktahuan akan rumus dasar menunjukkan bahwa proses pembelajaran mungkin terlalu berfokus pada aplikasi tanpa memastikan penguasaan konsep fundamental terlebih dahulu (Lestari et al., 2025; Sholichah & Rahayuningsih, 2025).

Kesalahan sistematis kedua, dan mungkin paling krusial karena dampak berantainya, ditemukan pada langkah perhitungan komponen *Sum of Squares* (SS), yaitu SS_{xy} , SS_{xx} , dan SS_{yy} . Akar masalah dari kesalahan ini bersifat konseptual, di mana siswa secara konsisten gagal membedakan antara notasi $\sum X^2$ (jumlah dari data yang dikuadratkan) dengan $(\sum X)^2$ (kuadrat dari total jumlah data). Miskonsepsi ini, seperti yang terekam dalam Gambar 4, menyebabkan kesalahan fatal dalam perhitungan, bahkan menghasilkan nilai SS negatif yang secara statistik tidak logis. Kesalahan pada tahap awal ini secara otomatis menjamin bahwa nilai r dan r^2 yang dihitung selanjutnya akan salah, menunjukkan betapa rapuhnya pemahaman siswa terhadap fondasi matematis dari analisis korelas (Luu-Thi et al., 2021; Razak et al., 2018)i.

Kesalahan paling umum ketiga teridentifikasi pada tahap akhir analisis, yaitu interpretasi data, khususnya dalam memahami makna koefisien determinasi (r^2). Banyak siswa tidak menyadari bahwa r^2 merupakan turunan langsung dari r , sehingga kesalahan perhitungan r secara otomatis menghasilkan interpretasi r^2 yang juga salah. Lebih lanjut, seperti yang ditunjukkan pada Gambar 5, siswa yang berhasil menghitung nilai cenderung memberikan interpretasi yang tidak lengkap, yakni tanpa menyebutkan persentase kontribusi variabel bebas terhadap variabel terikat. Selain itu, muncul pula kekeliruan dalam memaknai hubungan, seperti menyatakannya sebagai hubungan kausalitas (sebab-akibat), padahal analisis korelasi hanya menunjukkan kekuatan dan arah hubungan *linier* semata (Hung et al., 2017; Janse et al., 2021; Ranganathan & Aggarwal, 2016; Shao et al., 2022).

Secara kolektif, pola-pola kesalahan yang teridentifikasi—mulai dari ketidaktahuan rumus, miskonsepsi perhitungan SS , hingga kegagalan interpretasi r^2 —menunjukkan adanya disonansi yang jelas antara pengajaran prosedural dan pemahaman konseptual. Siswa mungkin diajarkan untuk "menghitung" tetapi tidak "memahami". Tingginya angka kesalahan (50%) menyiratkan bahwa metode pembelajaran yang ada saat ini mungkin belum cukup efektif dalam membangun literasi statistik yang utuh. Penekanan yang berlebihan pada hafalan rumus, tanpa pemahaman mendalam tentang logika di balik setiap komponen rumus, terbukti gagal membekali siswa dengan kompetensi yang diperlukan untuk menerapkan statistika secara benar.

Implikasi pedagogis dari temuan ini menunjukkan bahwa guru sebaiknya merancang ulang metode pengajaran korelasi agar tidak hanya menekankan latihan prosedural, melainkan juga mengutamakan pemahaman konsep secara mendalam. Contohnya, untuk membantu siswa mengatasi miskonsepsi pada perhitungan SS seperti yang terlihat di Gambar 4, guru perlu menjelaskan secara rinci perbedaan antara notasi $\sum X^2$ dan $(\sum X)^2$, menggunakan data sederhana agar lebih mudah dipahami. Selain itu, guna memperbaiki kelemahan dalam interpretasi hasil terutama pada r^2 yang ditunjukkan di Gambar 5 pengajaran harus menekankan bahwa siswa wajib menuliskan kesimpulan dalam bentuk kalimat lengkap, misalnya dengan menjelaskan

persentase kontribusi r^2 dalam konteks nyata seperti "persentase kontribusi jumlah kendaraan bermotor terhadap tingkat polusi udara". Dengan cara ini, fokus pembelajaran berpindah dari sekadar angka ke pemaknaan yang lebih relevan dan kontekstual bagi siswa.

Penelitian ini berhasil memetakan secara rinci tipologi dan frekuensi kesalahan siswa dalam analisis korelasi *Pearson* di satu sekolah spesifik (Ferwinda & Syahriluddin, 2019; Permata et al., 2025; Razak et al., 2018). Temuan ini memberikan data diagnostik yang berharga bagi guru untuk merancang intervensi pembelajaran yang lebih tepat sasaran. Namun, penelitian ini memiliki keterbatasan, yakni hanya dilakukan di SMA Negeri 3 Sidoarjo, sehingga temuan ini mungkin tidak dapat digeneralisasi ke populasi yang lebih luas tanpa kehati-hatian. Oleh karena itu, penelitian di masa depan disarankan untuk mereplikasi studi ini di konteks sekolah yang berbeda. Selain itu, sebuah studi eksperimental dapat dirancang untuk menguji efektivitas intervensi pengajaran yang secara spesifik menargetkan miskonsepsi (seperti pemahaman SS dan r^2) yang telah diidentifikasi dalam penelitian ini.

KESIMPULAN

Penelitian ini secara konklusif membuktikan bahwa siswa kelas XI di SMA Negeri 3 Sidoarjo mengalami kesulitan yang kompleks pada materi analisis korelasi. Kesulitan tersebut teridentifikasi pada dua level, yaitu prosedural dan konseptual. Dari sisi prosedural, mayoritas siswa (18 dari 36 orang) tidak mampu menghitung koefisien korelasi *Pearson* (r) dengan benar, yang umumnya diakibatkan oleh kesalahan mendasar dalam perhitungan SS_{xy} , SS_{xx} , dan SS_{yy} , baik karena penerapan rumus yang salah maupun kesalahan numerik dasar. Namun, permasalahan yang lebih mendalam terjadi pada level konseptual. Banyak siswa juga gagal dalam memberikan interpretasi yang sesuai terhadap nilai koefisien determinasi (r^2). Kegagalan ini—yang sering diperburuk oleh kesalahan kalkulasi sebelumnya—menunjukkan kelemahan mendasar pada pemahaman makna statistik di balik angka yang dihasilkan. Implikasi utama dari temuan ini ialah bahwa pembelajaran statistik di SMA perlu mengalami pergeseran, tidak semata berfokus pada hafalan rumus dan kemampuan kalkulasi, namun lebih menekankan pemahaman makna serta interpretasi data dalam konteks yang relevan. Penelitian lanjutan sangat disarankan untuk berpindah dari tahap penelusuran kesalahan menuju pengujian efektivitas solusi intervensi seperti model pembelajaran kontekstual atau *problem-based learning* untuk meningkatkan pemahaman konsep korelasi. Studi kualitatif berbasis wawancara kognitif juga perlu dilakukan guna menggali lebih rinci proses berpikir dan titik miskonsepsi siswa secara spesifik, sehingga membantu pengembangan strategi remedial yang lebih efektif

DAFTAR PUSTAKA

- Aswandari, A., Maraharani, S. D., & Susanti, R. (2025). Analisis kebutuhan pengembangan media flashcard berbasis kearifan lokal Musi Banyuasin sebagai alat bantu pembelajaran penjumlahan di kelas I sekolah dasar. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 5(2), 680. <https://doi.org/10.51878/science.v5i2.5221>
- Boyi, M. A., & Rahayuningsih, S. (2025). Analisis berpikir kritis siswa dalam mengerjakan soal turunan fungsi aljabar. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 5(3), 1266. <https://doi.org/10.51878/science.v5i3.6672>
- Damayanti, D. (2024). Analisis kesalahan representasi matematis siswa pada materi sistem persamaan linear dua variabel berdasarkan teori Nolting. *Cendekia Jurnal Ilmu Pengetahuan*, 4(4), 432. <https://doi.org/10.51878/cendekia.v4i4.3365>
- Damayanti, I. M., & Muslim, A. (2025). Peningkatan prestasi belajar dan sikap gotong-royong IPAS melalui model Auditory Intellectually Repetition (AIR) dengan

- menggunakan media diorama. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 5(3), 1148. <https://doi.org/10.51878/science.v5i3.6073>
- Ferwinda, E., & Syahriluddin, S. (2019). Analisis kesalahan siswa dalam mengerjakan soal cerita matematika berdasarkan tipe Newman pada kelas IV SD se-Gugus 1 Kecamatan Marpoyan Damai Pekanbaru. *Jurnal Pajar (Pendidikan Dan Pengajaran)*, 3(2). <https://doi.org/10.33578/pjr.v3i2.6664>
- Gao, J. (2023). R-squared (R2) – How much variation is explained? *Research Methods In Medicine & Health Sciences*, 5(4), 104. <https://doi.org/10.1177/26320843231186398>
- Hung, M., Bounsga, J., & Voss, M. W. (2017). Interpretation of correlations in clinical research. *Postgraduate Medicine*, 129(8), 902. <https://doi.org/10.1080/00325481.2017.1383820>
- Janse, R. J., Hoekstra, T., Jager, K. J., Zoccali, C., Tripepi, G., Dekker, F. W., & Diepen, M. van. (2021). Conducting correlation analysis: Important limitations and pitfalls. *Clinical Kidney Journal*, 14(11), 2332. <https://doi.org/10.1093/ckj/sfab085>
- Jumrah, M., Putri, R. M., & Afriani, D. (2023). Faktor-faktor kesulitan belajar matematika pada siswa SMA. *Jurnal Pendidikan Dan Pembelajaran Matematika*, 11(1), 67–75. <https://ejournal.unib.ac.id/index.php/jppm/article/view/26754>
- Kusumaningpuri, W., Andriyani, R., & Siregar, L. (2022). Statistika dalam pengambilan keputusan berbasis data. *Jurnal Matematika Dan Aplikasinya*, 10(2), 93–101. <https://ejournal.umm.ac.id/index.php/jmat/article/view/21032>
- Lantang, M., Nurhaliza, A., & Wulandari, D. (2021). Analisis kesalahan siswa dalam menyelesaikan soal matematika berdasarkan teori Watson. *Jurnal Pendidikan Matematika*, 9(1), 32–40. <https://ejurnal.untag-smd.ac.id/index.php/jpm/article/view/10119>
- Lestari, M. I., Lusiana, L., & Wahyuningsih, S. (2025). Analisis kesulitan belajar matematika berbasis masalah pada materi operasi hitung perkalian dan pembagian. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 5(3), 1285. <https://doi.org/10.51878/science.v5i3.6670>
- Luu-Thi, H.-T., Ngo-Thi, T.-T., Nguyen-Thi, M.-T., Ly, T.-T., Nguyen-Duong, B.-T., & Tran-Chi, V.-L. (2021). An investigation of mathematics anxiety and academic coping strategies among high school students in Vietnam: A cross-sectional study. *Frontiers In Education*, 6. <https://doi.org/10.3389/feduc.2021.742130>
- Permata, J. I., Budiarto, M. T., & Fuad, Y. (2025). Analysis of error in mathematical modeling problems: An exploratory study of middle school students. *Edelweiss Applied Science And Technology*, 9(7), 947. <https://doi.org/10.55214/25768484.v9i7.8773>
- Ranganathan, P., & Aggarwal, R. (2016). Common pitfalls in statistical analysis: The use of correlation techniques. *Perspectives In Clinical Research*, 7(4), 187. <https://doi.org/10.4103/2229-3485.192046>
- Razak, F. A., Baharun, N., Deraman, N. A., & Ismail, N. (2018). Assessing students' abilities in interpreting the correlation and regression analysis. *Journal Of Fundamental And Applied Sciences*, 9, 644. <https://doi.org/10.4314/jfas.v9i5s.45>
- Rizka, R. S. P., Sari, D. K., & Martusyilia, R. (2025). Penerapan pembelajaran berdiferensiasi dengan model Problem Based Learning dalam meningkatkan hasil belajar siswa. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 5(3), 1372. <https://doi.org/10.51878/science.v5i3.5625>
- Shao, K., Shirvan, M. E., & Alamer, A. (2022). How accurate is your correlation? Different methods derive different results and different interpretations. *Frontiers In*

Psychology, 13. <https://doi.org/10.3389/fpsyg.2022.901412>

Sholichah, M., & Rahayuningsih, S. (2025). Implementasi teknik scaffolding dalam pembelajaran matematika di SMA Negeri 1 Balen. *Learning Jurnal Inovasi Penelitian Pendidikan Dan Pembelajaran*, 5(3), 1529. <https://doi.org/10.51878/learning.v5i3.6115>

Tobing, S., Dharma, S., Mikael, S., Panjaitan, H., & Pakpahan, R. (2025). Pengaruh penggunaan video animasi pada mata pelajaran pendidikan Pancasila untuk meningkatkan motivasi belajar siswa kelas VII SMP Negeri 4 Tarutung. *Social Jurnal Inovasi Pendidikan IPS*, 5(3), 1133. <https://doi.org/10.51878/social.v5i3.6907>

Triyani, E., & Yuhana, M. (2023). Analisis kesalahan berdasarkan teori Watson pada penyelesaian soal matematika. *Jurnal Pendidikan Matematika*, 14(1), 51–59. <https://ejournal.upgris.ac.id/index.php/pendidikanmatematika/article/view/9512>

Winarto, B., Nuraina, N., Damayanti, N. W., & WIWI, Y. (2024). Analisis miskONSEPsiswa sekolah menengah pada telaah soal serta penggerjaannya. *Science Jurnal Inovasi Pendidikan Matematika Dan IPA*, 4(3), 176. <https://doi.org/10.51878/science.v4i3.3130>