

Penerapan Data Science Untuk Mengubah Data Menjadi Informasi Dan Pengetahuan Melalui Proses Pengumpulan, Pembersihan, Pengolahan, Dan Analisis Data Bagi Siswa SMKN 1 Warunggunung

Joko Yuwono^{1*}, Ita Hardianti², Muhsin Zaky Ramadhan³, Muhamad Farid Al Fachreza⁴, Zaskia Artika Rosadi⁵, Hanifa Sultoni⁶, Adam Rahmatullah⁷

^{1,2,3,4,5,6,7}Universitas Pamulang Kampus Kota Serang, Sistem Informasi, Universitas Pamulang, Serang, Indonesia

Email: ¹dosen02929@unpam.ac.id, ²dosen03355@unpam.ac.id, ³blackglory59@gmail.com, ⁴muhamadfaridalfachreza@gmail.com, ⁵zaskiaartika25@gmail.com, ⁶hanifasultoni8@gmail.com, ⁷adamr3540@gmail.com

*Email Corresponding Author: dosen02929@unpam.ac.id

Abstrak

Literasi data merupakan kompetensi penting bagi siswa SMK, tetapi siswa mitra belum memahami alur pengolahan data dari pengumpulan hingga analisis. Pengabdian ini bertujuan meningkatkan pemahaman dan literasi data siswa melalui penerapan proses data science. Kegiatan berupa workshop satu hari (125 menit) dengan pendekatan *learning by doing* di SMKN 1 Warunggunung, Kabupaten Lebak, Banten, diikuti 32 siswa kelas XI program keahlian Desain Komunikasi Visual. Materi mencakup pengumpulan, pembersihan, pengolahan, dan analisis data, serta penyajian tabel dan grafik menggunakan Microsoft Excel atau Google Sheets. Evaluasi memakai pre-test dan post-test lima butir soal serta observasi. Nilai rata-rata meningkat dari 38,13 menjadi 80,63 (naik 111,48%) dengan N-Gain 0,69 (kategori sedang). Seluruh peserta mengalami peningkatan, dan siswa berkategori Kurang turun dari 81,3% menjadi 0%. Ketepatan jawaban tentang urutan tahapan data science naik dari 18,75% menjadi 84,38%, sedangkan penerapan pada kasus desain baru mencapai 68,75% sehingga perlu diperkuat.

Kata Kunci: data science, literasi data, pembersihan data, spreadsheet, siswa SMK

Abstract

Data literacy is an essential competency for vocational high school students, yet the partner school's students did not fully understand the data-processing flow from collection to analysis. This service aimed to improve students' data literacy through the data science process. A one-day workshop (125 minutes) using a learning-by-doing approach was held at SMKN 1 Warunggunung, Lebak Regency, with 32 eleventh-grade students majoring in Visual Communication Design. The material covered data collection, cleaning, processing, and analysis, as well as presenting tables and charts using Microsoft Excel or Google Sheets. Evaluation used a five-item pre-test, post-test, and observation. The mean score rose from 38.13 to 80.63 (an increase of 111.48%) with an N-Gain of 0.69 (medium category). All participants improved, and the share of students in the lowest category fell from 81.3% to 0%. Correct answers on the sequence of data science stages rose from 18.75% to 84.38%, whereas application to design cases reached only 68.75% and needs reinforcement.

Keywords: data science, data literacy, data cleaning, spreadsheet, vocational students

1. PENDAHULUAN

Perkembangan teknologi informasi mendorong pertumbuhan volume data di hampir semua bidang kehidupan. Data mentah baru bernilai guna setelah diolah secara sistematis menjadi informasi dan pengetahuan (Marchionini, 2023; Wolski & Gomolińska, 2020). Data science memadukan statistika, ilmu komputer, dan pengetahuan domain untuk mengekstraksi makna dari data, dan kemampuan ini semakin dibutuhkan pada era Industri 4.0 (Sasongko, 2023; Syamsu & Widodo, 2021). Proses kerjanya umumnya terdiri atas empat tahap, yaitu pengumpulan, pembersihan, pengolahan, dan analisis data.

Mitra kegiatan ini adalah SMKN 1 Warunggunung, sekolah menengah kejuruan negeri di Jalan Raya Petir Km. 05, Desa Jagabaya, Kecamatan Warunggunung, Kabupaten Lebak, Provinsi Banten. Sekolah ini menyelenggarakan empat kompetensi keahlian, yaitu Bisnis Retail, Desain Komunikasi Visual (DKV), Agribisnis Perikanan Air Tawar, dan Agribisnis Pengolahan Hasil Perikanan, dengan orientasi pada kesiapan siswa memasuki dunia usaha dan dunia industri. Sasaran kegiatan adalah 32 siswa kelas XI program keahlian DKV. Potret kemampuan awal mereka rendah: rata-rata pre-test hanya 38,13 dari skala 100, 26 siswa (81,3%) berada pada kategori Kurang, dan tidak ada siswa pada kategori Baik. Hanya 6 siswa (18,75%) yang dapat menjawab benar urutan tahapan data science, dan 12 siswa (37,50%) memahami tujuan pembersihan data. Padahal, siswa DKV perlu membaca data dalam pekerjaannya, misalnya untuk menganalisis tren peminatan perangkat lunak desain atau nilai tugas.

Sejumlah kegiatan pengabdian melaporkan bahwa pelatihan data science dan literasi data pada siswa sekolah menengah dapat meningkatkan pemahaman. Khairi et al. (2024) melatih siswa SMKN 2 Kraksaan memakai Python, dengan kenaikan skor 20–30 poin dan sekitar 70% siswa menyelesaikan tugas klasifikasi. Ajiningrum et al. (2025) memakai Excel dan Google Sheets pada data biologi untuk siswa SMA dan melaporkan kenaikan pemahaman rata-rata 37,3%. Afitriansyah et al. (2025) melaporkan kenaikan pemahaman konsep *big data* sebesar 41,6% pada siswa SMK Teknik Komputer dan Jaringan setelah pelatihan Excel, sedangkan Dewi dan Dewi (2026) mencatat kenaikan skor rata-rata 71,7% pada pelatihan visualisasi data berbasis Excel untuk siswa SMK akuntansi. Putri et al. (2024) memperkenalkan data science lewat simulasi permainan dan menemukan peningkatan motivasi terhadap STEM, Eliyani et al. (2024) menerapkan *clustering* untuk mengembangkan keterampilan analitik siswa SMK, dan Merdikawati et al. (2026) melaporkan kenaikan pengetahuan 27% pada lokakarya data science untuk masyarakat umum. Sepanjang penelusuran penulis, kegiatan-kegiatan tersebut cenderung berfokus pada satu teknik atau satu bidang, dan belum banyak yang melaporkan capaian per butir soal dari alur lengkap pengumpulan sampai analisis data pada siswa DKV.

Berdasarkan kondisi tersebut, masalah yang dikaji adalah (1) bagaimana alur data science yang mencakup pengumpulan, pembersihan, pengolahan, dan analisis data dapat diterapkan pada siswa SMK dengan perangkat yang mudah diakses, dan (2) seberapa besar peningkatan pemahaman dan literasi data siswa setelah mengikuti kegiatan tersebut. Tujuan pengabdian ini adalah melatih siswa menjalankan empat tahap data science memakai spreadsheet dan mengukur peningkatan pemahaman mereka.

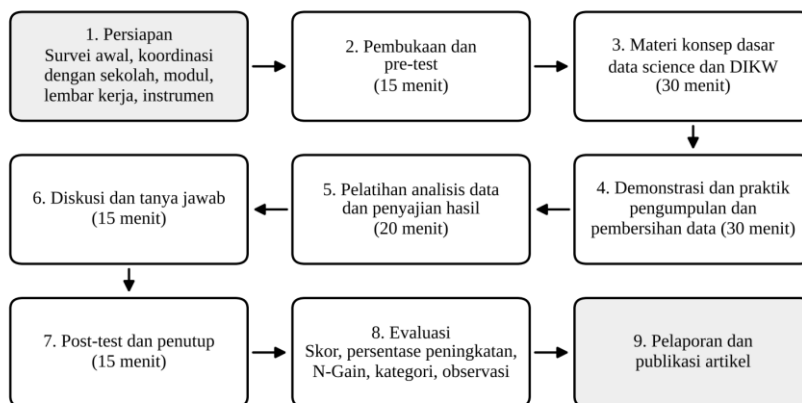
2. METODE PELAKSANAAN

Kegiatan dilaksanakan pada 24 September 2026 di SMKN 1 Warunggunung dengan 32 siswa kelas XI program keahlian DKV sebagai peserta. Metode yang digunakan adalah workshop satu hari berdurasi 125 menit dengan pendekatan *learning by doing*, yang menggabungkan ceramah interaktif, demonstrasi, praktik langsung, diskusi dan tanya jawab, serta penyusunan laporan analisis sederhana. Perangkat yang dipakai adalah Microsoft Excel atau Google Sheets, dengan data latihan yang dekat dengan lingkungan sekolah, seperti data nilai, minat belajar, dan kegiatan ekstrakurikuler. Pelaksanaan terdiri atas tiga tahap. Tahap persiapan meliputi survei awal ke sekolah, koordinasi jadwal dan peserta, serta penyusunan materi, lembar kerja, data latihan, dan instrumen evaluasi. Tahap pelaksanaan mengikuti urutan pada Tabel 1 dan Gambar 1. Tahap evaluasi dan pelaporan meliputi pengolahan skor, observasi, penyusunan laporan, dan publikasi artikel.

Tabel 1. Tahapan pelaksanaan kegiatan

Tahap	Kegiatan	Durasi
1	Pembukaan dan pre-test	15 menit
2	Penyampaian materi konsep dasar data science dan hierarki DIKW	30 menit
3	Demonstrasi dan praktik pengumpulan data dan pembersihan data	30 menit
4	Pelatihan analisis data dan penyajian hasil dalam tabel dan grafik	20 menit

Tahap	Kegiatan	Durasi
5	Diskusi dan tanya jawab	15 menit
6	Post-test dan penutup	15 menit
Total		125 menit



Gambar 1. Diagram alur pelaksanaan kegiatan pengabdian

Keberhasilan kegiatan diukur dengan pre-test dan post-test berisi lima butir soal pilihan ganda yang sama, dengan bobot 20 poin untuk setiap jawaban benar sehingga skor berkisar 0–100. Kelima butir mengukur pemahaman tentang (Q1) tujuan utama pembersihan data, (Q2) urutan tahapan data science, (Q3) alasan data mentah tidak langsung dapat dianalisis, (Q4) perubahan data mentah menjadi informasi dan pengetahuan setelah dianalisis, dan (Q5) contoh penerapan sederhana data science bagi siswa DKV. Data dianalisis secara deskriptif melalui rata-rata, median, simpangan baku, persentase ketepatan jawaban per butir, dan persentase peningkatan, yang dihitung sebagai $(\text{rata-rata post-test} - \text{rata-rata pre-test}) \div \text{rata-rata pre-test} \times 100$. Besarnya peningkatan juga dinyatakan dengan N-Gain, yaitu $g = (\text{skor post-test} - \text{skor pre-test}) \div (100 - \text{skor pre-test})$, dengan kategori tinggi ($g \geq 0,70$), sedang ($0,30 \leq g < 0,70$), dan rendah ($g < 0,30$) (Hake, 1998). Skor siswa juga dikelompokkan menjadi kategori Sangat Baik (86–100), Baik (71–85), Cukup (56–70), dan Kurang (≤ 55).

Perubahan sikap peserta dinilai secara deskriptif-kualitatif melalui observasi keterlibatan selama praktik dan diskusi, yaitu keaktifan bertanya, kemampuan menyelesaikan latihan, dan kesediaan mempresentasikan hasil. Dampak pada aspek sosial budaya dan ekonomi tidak diukur secara langsung karena kegiatan berlangsung satu hari. Aspek tersebut dibahas sebagai dampak jangka panjang berupa kesiapan siswa menghadapi dunia kerja.

3. HASIL PEMBAHASAN

3.1. Pelaksanaan Kegiatan

Pada sesi pertama, peserta diperkenalkan dengan hierarki data, informasi, pengetahuan, dan kebijaksanaan (DIKW) serta peran data science di era Industri 4.0. Peserta diajak memahami bahwa data mentah belum bernilai guna sebelum diolah secara sistematis (Marchionini, 2023; Syamsu & Widodo, 2021). Pada sesi berikutnya, instruktur mendemonstrasikan penghimpunan data primer dan sekunder serta cara mengenali data hilang (*missing value*), data duplikat, dan format yang tidak seragam di Excel atau Google Sheets, kemudian peserta mengulang langkah yang sama pada data latihan dengan pendampingan. Pada sesi pengolahan dan analisis, peserta mengelompokkan, mengurutkan, dan mentransformasi nilai, melakukan analisis deskriptif sederhana, lalu menyajikan hasilnya dalam tabel dan grafik. Beberapa peserta mempresentasikan hasil kerjanya sehingga terjadi pertukaran temuan.

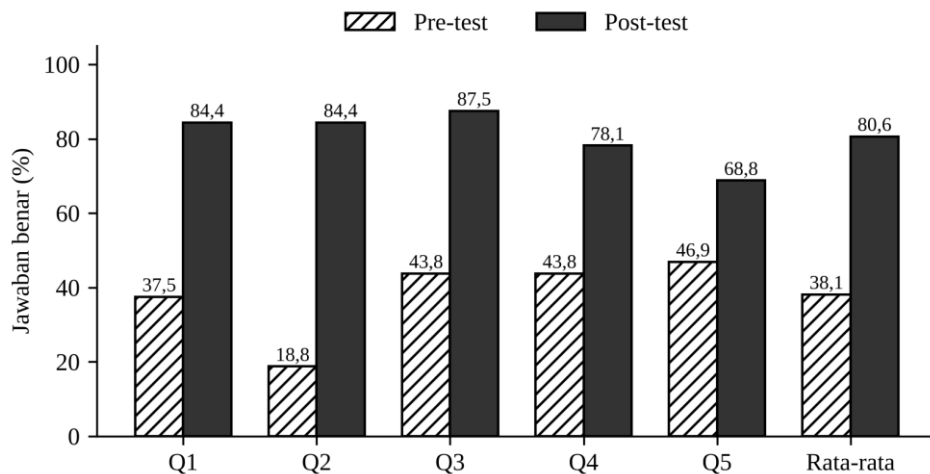
Sesi diskusi berlangsung interaktif. Pertanyaan peserta berkisar pada (1) cara menentukan data yang layak dibersihkan atau dihapus, (2) pemilihan jenis grafik yang sesuai dengan data, dan (3) pemanfaatan data science pada bidang keahlian mereka, seperti bisnis retail, desain komunikasi visual, dan agribisnis perikanan. Pertanyaan ketiga menunjukkan bahwa peserta mulai mengaitkan data dengan bidang keahliannya, yang merupakan perubahan sikap yang diharapkan.

3.2. Peningkatan Pemahaman dan Literasi Data

Ketepatan jawaban meningkat pada kelima butir soal (Tabel 2 dan Gambar 2). Peningkatan terbesar terjadi pada Q2 tentang urutan tahapan data science, dari 18,75% (6 siswa) menjadi 84,38% (27 siswa), naik 65,63 poin persentase. Butir ini merupakan yang tersulit pada pre-test. Peningkatan besar juga terjadi pada Q1 tentang tujuan pembersihan data (naik 46,88 poin) dan Q3 tentang alasan data mentah tidak langsung dapat dianalisis (naik 43,75 poin), dengan ketepatan post-test tertinggi pada Q3, yaitu 87,50%. Peningkatan terendah terdapat pada Q5 tentang contoh penerapan pada kasus DKV, dari 46,88% menjadi 68,75% (naik 21,88 poin). Peserta lebih mudah memahami konsep dan urutan tahapan daripada menerapkannya pada kasus di bidang keahlian mereka.

Tabel 2. Rekapitulasi ketepatan jawaban per butir soal (N = 32)

Butir	Indikator	Pre-test n (%)	Post-test n (%)	Selisih (poin %)
Q1	Tujuan pembersihan data	12 (37,50)	27 (84,38)	+46,88
Q2	Urutan tahapan data science	6 (18,75)	27 (84,38)	+65,63
Q3	Alasan data mentah tidak langsung dianalisis	14 (43,75)	28 (87,50)	+43,75
Q4	Perubahan data menjadi informasi dan pengetahuan	14 (43,75)	25 (78,13)	+34,38
Q5	Penerapan data science bagi siswa DKV	15 (46,88)	22 (68,75)	+21,88
Rata-rata		12,2 (38,13)	25,8 (80,63)	+42,50



Gambar 2. Persentase jawaban benar per butir soal pada pre-test dan post-test

Nilai rata-rata siswa naik dari 38,13 pada pre-test menjadi 80,63 pada post-test, yaitu naik 42,50 poin atau 111,48% (Tabel 3). N-Gain kelompok sebesar 0,69 termasuk kategori sedang dan mendekati batas kategori tinggi, sedangkan rata-rata N-Gain individu sebesar 0,71. Seluruh siswa mengalami peningkatan, dengan kenaikan terkecil

20 poin dan terbesar 60 poin. Berdasarkan kategori N-Gain individu, 10 siswa (31,3%) berkategori tinggi dan 22 siswa (68,8%) berkategori sedang, tanpa siswa berkategori rendah.

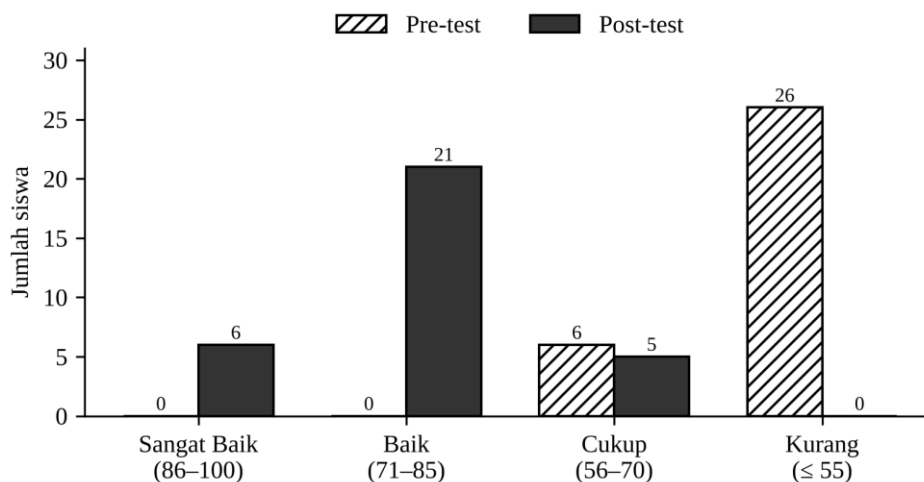
Tabel 3. Ringkasan statistik deskriptif pre-test dan post-test (N = 32)

Ukuran	Pre-test	Post-test	Selisih / capaian
Nilai rata-rata	38,13	80,63	+42,50 poin (111,48%)
Median	40	80	+40 poin
Nilai tertinggi	60	100	+40 poin
Nilai terendah	20	60	+40 poin
Simpangan baku	13,78	11,90	–
N-Gain kelompok	–	–	0,69 (sedang)
Rata-rata N-Gain individu	–	–	0,71
Siswa yang nilainya meningkat	–	–	32 dari 32 siswa (100%)

Distribusi kategori nilai bergeser ke arah yang lebih baik (Tabel 4 dan Gambar 3). Pada pre-test, 26 siswa (81,3%) berada pada kategori Kurang dan 6 siswa (18,8%) pada kategori Cukup. Setelah kegiatan, tidak ada lagi siswa pada kategori Kurang, 21 siswa (65,6%) berada pada kategori Baik, 6 siswa (18,8%) pada kategori Sangat Baik, dan 5 siswa (15,6%) pada kategori Cukup. Kelima siswa pada kategori Cukup tersebut memerlukan pendampingan lanjutan.

Tabel 4. Distribusi kategori nilai pre-test dan post-test

Kategori	Rentang	Pre-test n (%)	Post-test n (%)
Sangat Baik	86–100	0 (0,0)	6 (18,8)
Baik	71–85	0 (0,0)	21 (65,6)
Cukup	56–70	6 (18,8)	5 (15,6)
Kurang	≤ 55	26 (81,3)	0 (0,0)
Jumlah		32 (100)	32 (100)



Gambar 3. Jumlah siswa menurut kategori nilai pada pre-test dan post-test

3.3. Indikator Ketercapaian dan Tolak Ukur Keberhasilan

Ketercapaian dinilai dengan empat indikator. Pertama, peningkatan skor: rata-rata naik 111,48% dengan N-Gain 0,69 yang termasuk kategori sedang menurut Hake (1998). Kedua, pergeseran kategori nilai: siswa berkategori Kurang turun dari 81,3% menjadi 0%, dan 27 dari 32 siswa (84,4%) mencapai kategori Baik atau Sangat Baik. Ketiga, sebaran peningkatan: seluruh peserta (100%) memperoleh skor lebih tinggi, dengan kenaikan minimal 20 poin. Keempat, keterlibatan peserta: peserta menyelesaikan latihan dan bertanya serta mempresentasikan hasil kerjanya. Indikator yang belum optimal adalah penerapan pada kasus DKV (Q5) dengan ketepatan post-test 68,75%. Hasil ini sejalan dengan kegiatan serupa yang melaporkan peningkatan pemahaman setelah pelatihan data science berbasis praktik (Afitriansyah et al., 2025; Ajiningrum et al., 2025; Khairi et al., 2024). Kenaikan relatif pada kegiatan ini (111,48%) lebih besar daripada yang dilaporkan Afitriansyah et al. (2025) sebesar 41,6% dan Dewi dan Dewi (2026) sebesar 71,7%, tetapi angka tersebut tidak dapat dibandingkan langsung karena skor awal peserta di sini rendah (38,13), sedangkan instrumen dan jumlah butir soal berbeda.

3.4. Keunggulan dan Kelemahan Kegiatan

Keunggulan kegiatan ini adalah: (1) berbasis praktik, sehingga setiap peserta menjalankan sendiri tahap data science pada perangkatnya, sesuai dengan karakter siswa SMK yang menyukai pembelajaran terapan; (2) mencakup empat tahap data science sekaligus, bukan satu teknik tunggal; (3) memakai Excel dan Google Sheets yang sudah dikenal siswa dan tidak memerlukan instalasi atau spesifikasi perangkat tinggi; (4) memakai data latihan yang kontekstual; dan (5) memiliki evaluasi kuantitatif pada hari yang sama sehingga peningkatan dapat diukur. Kelemahannya adalah: (1) desain satu kelompok tanpa kelompok pembandingan; (2) post-test diberikan pada hari yang sama dengan soal yang sama, sehingga hasilnya menggambarkan pemahaman jangka pendek dan dapat dipengaruhi kesamaan soal; (3) instrumen hanya lima butir sehingga skor bervariasi dalam kelipatan 20 dan ketelitian pengukuran terbatas; dan (4) retensi pengetahuan belum diuji.

3.5. Kendala dan Strategi Penyelesaian

Tingkat kesulitan pelaksanaan tergolong sedang. Kendala utama berasal dari keragaman kemampuan awal peserta dan keterbatasan waktu, bukan dari kompleksitas materi. Kendala dan strategi penyelesaiannya disajikan pada Tabel 5.

Tabel 5. Kendala dan strategi penyelesaian dalam pelaksanaan kegiatan

No	Kendala	Strategi penyelesaian
1	Kemampuan awal beragam; sebagian siswa belum mengenal fungsi dasar seperti <i>sort</i> dan <i>filter</i>	Pendampingan oleh tim, <i>peer tutoring</i> oleh siswa yang lebih mahir, dan latihan yang tingkat kesulitannya dinaikkan bertahap
2	Waktu terbatas pada satu hari	Materi dikemas ringkas pada konsep kunci; lembar kerja dan panduan praktik dibagikan untuk belajar mandiri
3	Tidak semua peserta memiliki laptop dan jaringan internet kurang stabil	Peserta dikelompokkan dan berbagi perangkat; data latihan disiapkan dalam berkas yang dapat dipakai tanpa internet
4	Istilah teknis (<i>data cleaning</i> , <i>missing value</i> , <i>data processing</i>) masih asing	Istilah dijelaskan dengan bahasa sederhana, contoh sehari-hari, dan padanan istilah dalam bahasa Indonesia
5	Peserta belum terbiasa menarik kesimpulan dari data	Instruktur memberi pertanyaan pemandu dan sesi diskusi untuk melatih interpretasi hasil analisis

3.6. Dampak Jangka Panjang dan Peluang Pengembangan

Bagi siswa, kegiatan ini menumbuhkan kemampuan berpikir berbasis data yang dapat dipakai dalam tugas kejuruan, praktik kerja lapangan, dan persiapan memasuki dunia kerja. Bagi guru, lembar kerja dan panduan praktik berbasis spreadsheet yang diserahkan kepada sekolah dapat dipakai ulang sebagai bahan ajar. Bagi sekolah, literasi data dapat dikembangkan sebagai kegiatan ekstrakurikuler atau proyek lintas kompetensi keahlian, misalnya analisis data penjualan pada Bisnis Retail atau data produksi pada Agribisnis Perikanan. Peluang pengembangan ke depan mencakup penambahan studi kasus yang dekat dengan bidang DKV, pelatihan visualisasi data, pengenalan Python dasar seperti yang dilakukan Khairi et al. (2024), serta evaluasi dengan lebih banyak butir soal, kelompok pembeding, dan tes lanjutan untuk mengukur retensi.

4. KESIMPULAN

Workshop data science satu hari untuk 32 siswa kelas XI program keahlian DKV di SMKN 1 Warunggunung meningkatkan nilai rata-rata dari 38,13 menjadi 80,63 (naik 111,48%) dengan N-Gain 0,69 (kategori sedang). Seluruh peserta mengalami peningkatan, siswa berkategori Kurang berkurang dari 26 menjadi nol, dan 27 dari 32 siswa mencapai kategori Baik atau Sangat Baik. Pemahaman tentang urutan tahapan data science meningkat paling besar (dari 18,75% menjadi 84,38%), sedangkan penerapan pada kasus DKV meningkat paling kecil (dari 46,88% menjadi 68,75%). Hambatan berupa kemampuan awal yang beragam, waktu terbatas, keterbatasan perangkat dan jaringan, istilah teknis yang asing, dan kesenjangan literasi data dapat diatasi melalui pendampingan, *peer tutoring*, penyederhanaan istilah, serta penyediaan lembar kerja dan data latihan luring. Untuk keberlanjutan, disarankan agar kegiatan dibagi ke dalam beberapa pertemuan, studi kasus bidang DKV diperbanyak, dan siswa yang masih berkategori Cukup mendapat pendampingan lanjutan. Sekolah dapat memanfaatkan lembar kerja yang telah diserahkan serta menyiapkan laboratorium komputer dan jaringan yang memadai. Kegiatan berikutnya sebaiknya memakai instrumen dengan lebih banyak butir soal, kelompok pembeding, dan tes lanjutan untuk mengukur retensi, serta dapat memperluas materi ke visualisasi data dan Python dasar.

5. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Pamulang yang mendanai kegiatan ini melalui program Pengabdian kepada Masyarakat Tahun Akademik 2026/2027. Terima kasih juga disampaikan kepada pimpinan, guru, dan siswa SMKN 1 Warunggunung atas kerja sama dan partisipasinya.

6. REFERENSI

- Afitriansyah, H., Fathan, F. N., Muhammad, R. N., Utomo, H. P., & Firmansyah, H. S. (2025). Peningkatan literasi big data melalui pelatihan terapan bagi siswa SMK Teknik Komputer dan Jaringan. *JMM (Jurnal Masyarakat Mandiri)*, 9(4), 4471–4480. <https://doi.org/10.31764/jmm.v9i4.33330>
- Ajiningrum, P. S., Binawati, D. K., Kusuma, P. S. W., Sukarjati, Andriani, V., Sopandi, T., & Yachya, A. (2025). Pengenalan sains data biologi untuk siswa SMA: Membangun literasi data sejak dini dalam era digital. *Jurnal Abadimas Adi Buana*, 9(1), 95–102. <https://doi.org/10.36456/abadimas.v9.i01.a10561>
- Dewi, P. A. C., & Dewi, D. P. N. (2026). Pelatihan visualisasi data berbasis Excel untuk meningkatkan kemampuan literasi numerasi generasi Z. *JMM (Jurnal Masyarakat Mandiri)*, 10(2), 2366–2376. <https://doi.org/10.31764/jmm.v10i2.37734>
- Eliyani, E., Rifqi, M., & Dwiasnati, S. (2024). Penerapan clustering dalam data science untuk mengembangkan keterampilan analitik di SMK Media Informatika. *Jurnal Pengabdian Masyarakat*, 3(2), 218–228. <https://doi.org/10.70340/japamas.v3i2.162>

-
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Khairi, M., Riski, M., Rosiqin, Moh. M., Arifin, A. R., Irsyad, Moh. F. R., & Mardin. (2024). Peningkatan literasi data melalui pelatihan data science untuk siswa SMKN 2 Kraksaan. *Jurnal Pengabdian Masyarakat Sains dan Teknologi*, 3(4), 138–147. <https://doi.org/10.58169/jpmsaintek.v3i4.652>
- Marchionini, G. (2023). Information and data sciences: Context, units of analysis, meaning, and human impact. *Data and Information Management*, 7(1), 100031. <https://doi.org/10.1016/j.dim.2023.100031>
- Merdikawati, S., Herlambang, M. B., Sudri, N. M., Mauliddina, Y., & Ranti, G. (2026). Penguatan literasi digital melalui workshop data science dan machine learning berbasis pembelajaran interaktif. *Abdi Teknayasa*, 7(1). <https://journals2.ums.ac.id/abditeknayasa/article/view/13307>
- Putri, D. E., Baskara, Z. W., Harsyiah, L., Kurnia, A., Purnamasari, N. A., Hadijati, M., Hidayati, L., Andriani, H., Qudsi, J., Ilma, H., Asri, A. T. J. A., Lestari, Y., Melani, J., Fadhillah, R., Syahrul, M., & Husni, M. N. (2024). Pengenalan data science untuk mempersiapkan era digital pada siswa di SMAN 1 Gunung Sari. *Jurnal Pengabdian Magister Pendidikan IPA*, 7(4), 1189–1193. <https://doi.org/10.29303/jpmipi.v7i4.9022>
- Sasongko, A. T. (2023). Studi literatur konsep dan implementasi sains data untuk memaksimalkan kinerja industri manufaktur. *Jurnal Teknologi dan Sistem Informasi Bisnis*, 5(2), 90–94. <https://doi.org/10.47233/jteksis.v5i2.778>
- Syamsu, M., & Widodo, W. (2021). Peran data science dan data scientist untuk mentransformasi data dalam industri 4.0. *Jurnal Teknologi Informasi (JUTECH)*, 2(1), 27–36. <https://doi.org/10.32546/jutech.v2i1.1540>
- Wolski, M., & Gomolińska, A. (2020). Data meaning and knowledge discovery: Semantical aspects of information systems. *International Journal of Approximate Reasoning*, 119, 40–57. <https://doi.org/10.1016/j.ijar.2020.01.002>