

Workshop Peran Statistika Data Sains untuk Siswa-siswi SMA WASKITO**Workshop on the Role of Statistics in Science Data for Students of SMA WASKITO****Wiwik Wiyanti*, Dhela Asafiani Agatha, Ferdinand Nathaniel Widjaya**

Universitas Matana, Tangerang, Banten

*Email: wiwik.chae@gmail.com

(Diterima 10-07-2024; Disetujui 19-08-2024)

ABSTRAK

Permintaan data saintis yang mumpuni di Indonesia saat tulisan ilmiah ini dibuat sedang terus meningkat. Pengetahuan dasar statistika yang berhubungan dengan data sains tentunya diharapkan dapat dibekalkan kepada generasi muda sedini mungkin. Penelitian ini bertujuan untuk mengedukasi peran statistika dalam data sains kepada siswa-siswi SMA Waskito. Metode penelitian yang digunakan untuk keperluan analisa data adalah penelitian *Mixed-Method*. Adapun parameter kuantitatif yang diukur adalah apakah nilai rata-rata hasil pembelajaran melalui *workshop* peran statistika data sains untuk kelas XII SMA Waskito lebih dari 70 secara signifikan atau tidak. Sampel yang digunakan dalam penelitian ini adalah 176 siswa kelas XII SMA Waskito. Adapun analisa data menggunakan inferensia *One Mean T – test*. Dari hasil analisa data diperoleh kesimpulan bahwa dari data *post-test* yang diperoleh tidak cukup bukti untuk mengatakan bahwa rata-rata nilai *post-test* lebih besar dari 70. Adapun saran untuk PKM selanjutnya adalah mengelola kelas secara virtual ataupun menggunakan laboratorium untuk praktikum, supaya hasil pembelajaran efektif dan tepat guna.

Kata kunci: pengabdian, sains, data, statistika, *workshop***ABSTRACT**

Demand of qualified data scientist in Indonesia at this time was increasing. The basic knowledge of statistics related to data science is certainly expected to be provided to the younger generation as early as possible. This research aims to educate the role of statistics in data science to SMA Waskito students. The research method use Mixed-Method research. The quantitative parameters measure are whether the average of learning outcomes through the workshop is significantly more than 70 or not. The sample used in this research was 176 students of class XII SMA Waskito. The data analysis used One Mean T-test. From the results of the analysis, it was concluded that the post-test data obtained was not sufficient evidence to say that the average of post test was greater than 70 significantly.

Keywords: *pkm, science, data, statistics, workshop***PENDAHULUAN**

Abad 21 merupakan abad di mana inovasi teknologi berkembang sangat pesat. Sebagai contoh ketika wabah *covid* melanda dunia, membuat individu di dunia ini, khususnya masyarakat Indonesia menyesuaikan dengan perkembangan teknologi. Misalkan penyesuaian teknologi di bidang pendidikan. Dari siswa dan guru, yang semula sekolah tatap muka datang ke sekolah akhirnya dapat dijangkau dengan tatap muka via *online*. Di dunia bisnis, dari yang semula transaksi tatap muka di pasar, berubah menjadi lewat *e-commerce*. Dari sisi transportasi, ketika sebelumnya menggunakan ojek pangkalan, berubah memanfaatkan teknologi menjadi ojek *online*. Dari membeli tiket pesawat harus mencetak kertas, sekarang sudah memanfaatkan teknologi tanpa harus cetak kertas.

Dari contoh tersebut, merupakan peluang bagi data *scientist*/ilmuwan yang berperan penting dalam organisasi yang bergerak di bidang digital teknologi. Kenapa menjadi peluang

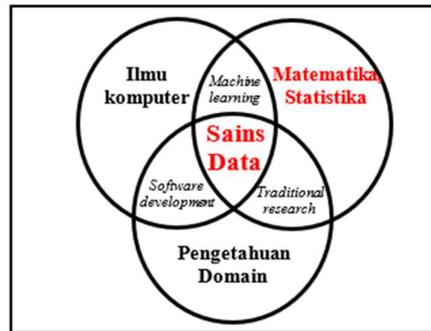
data *scientist*?. Dikutip tulisan Dirwan dalam laman web valianci.ai, alasannya karena di era revolusi industri 4.0 ini, berbagai sektor industri, publik, maupun nirlaba berlomba-lomba untuk memanfaatkan data yang tersedia di setiap organisasinya guna mendapatkan solusi, ide dan menentukan langkah organisasi selanjutnya. Bahkan analisa data merupakan faktor kunci untuk mengambil keputusan dalam bisnis. Seperti pada tahun 2012 Deloitte melakukan survei kepada perwakilan 75 perusahaannya yang tersebar di Kawasan Amerika Utara, Asia dan Inggris. Hasilnya adalah 49% responden menyatakan hal sama, bahwa analisa data merupakan faktor kunci untuk mengambil keputusan dalam bisnis (Ahdiat, 2023).

Perlu diketahui bahwa kebutuhan data saintis di Indonesia sendiri terus meningkat. Dalam laporan Mursid, 2012 di Indonesia, kebutuhan tenaga kerja praktisi data meningkat mencapai 9 juta praktisi data. Hal ini juga diperkuat tulisan Syafri dalam laman web ([unej](http://unej.ac.id), 2019) menurut Syafri Bahar, *Vice President* Data Gojek yang mengutip data Kominfo RI, setiap tahunnya Indonesia membutuhkan 600 ribu *data scientist* dari berbagai bidang seperti analis data, intelijen data, rekayasa data, kecerdasan buatan, matematikawan hingga aktuaris. Kebutuhan sebesar itu juga dapat dilihat dari laporan Forbes tahun 2020 ([dqlab](http://dqlab.com), 2020). Dikutip dalam laman web [dqlab](http://dqlab.com) berdasarkan laporan yang dikeluarkan oleh Forbes tahun 2020, profesi *Data Scientist* dan *data analytics* diperkirakan akan meningkat dari 364.000 lowongan menjadi 2.720.000. Namun, jika ditilik dari permintaan yang begitu besar, justru sampai tulisan ini dibuat, pemenuhan data saintis Indonesia yang mumpuni dan sesuai harapan pasar masih tergolong langka. Hal tersebut karena belum banyak sumber daya manusia Indonesia yang menguasai baik segi ilmu data sains maupun pengalaman (*experience*) yang telah diperoleh pelamar.

Oleh karena itu, pengetahuan mengenai data sains terus digencarkan, bahkan jika memungkinkan sejak dini. Seperti halnya, yang dilakukan oleh Kemkominfo, yang mengadakan pelatihan data sains untuk anak. Dikutip dari siaran pers Kemkominfo pada "(Peluncuran pelatihan) *Data Science for Kids* ini adalah satu bagian untuk bisa menyediakan sebanyak 50 ribu *data scientist* di seluruh Indonesia, jadi kita melakukan dari berbagai tingkat mulai tingkat dasar yaitu tingkat anak-anak, kemudian tingkat menengah sampai tingkat *advance*," kata Kepala Badan Penelitian dan Pengembangan Sumber Daya Manusia Kementerian Kominfo Hary Budiarto.

Perlu diketahui bahwa untuk menjadi data saintis, dibutuhkan keahlian yang dibangun dari tiga pilar ilmu, yaitu statistika, ilmu komputer serta pengetahuan dasar seperti bisnis, keuangan, kesehatan, dan lain sebagainya. Permasalahan mendasar, untuk mencetak data

santis muda yang tangguh, kiranya dasar ilmu data sains sudah harus diberikan dari bangku sekolah sebelum mematangkannya di bangku perkuliahan. Kenapa hal ini harus dilakukan? salah satunya adalah perkembangan era teknologi 4.0. Gambaran hubungan *skill* yang harus dikuasai data saintis, dapat dilihat pada gambar 1.



Gambar 1. Diagram *Venn* hubungan keilmuan pada data sains.

Sementara itu, temuan di lapangan, kurikulum pada sekolah kurikulum internasional sudah sangat jauh jika dibandingkan dengan kurikulum nasional di Indonesia. Berdasarkan temuan di lapangan, materi statistika kelas XII SMA Nasional mempelajari statistika sebatas statistika diskriptif, seperti menghitung *mean*, *median*, modus, histogram, membuat diagram dan menginterpretasikan (Achmad, 2020). Sementara itu, sekolah internasional sudah memberlakukan kurikulum statistika dengan taraf materi statistika dasar di bangku perkuliahan atau statistika inferensia. Alasan tersebut merupakan alasan mendasar yang menjadikan peneliti untuk melaksanakan pengabdian kepada masyarakat (PKM), yaitu *workshop* statistika data sains ke sekolah nasional SMK/SMA. Adapun sekolah yang dijadikan pelaksanaan PKM kali ini adalah sekolah menengah atas (SMA) Waskito, yang merupakan sekolah dengan kurikulum nasional.

Harapannya adanya *workshop* ini adalah menambah ilmu pengetahuan dasar statistika untuk data sains kepada siswa-siswi SMA Waskito. Lebih lanjut lagi adalah menjadikan iklim minat belajar statistika data sains yang selanjutnya adalah memilih jurusan statistika data sains, sehingga Indonesia akan menambah *data saintist* muda Indonesia yang mumpuni.

Berdasarkan uraian tersebut di atas, maka fokus PKM ini adalah untuk menganalisa apakah nilai rata-rata hasil pembelajaran melalui *workshop* peran statistika data sains untuk kelas XII SMA Waskito lebih dari 70 secara signifikan?. Hipotesa yang digunakan adalah:

H_0 : Nilai rata-rata pembelajaran melalui *workshop* peran statistika data sains untuk kelas XII SMA Waskito tidak lebih dari 70 secara signifikan.

H_a : Nilai rata-rata pembelajaran melalui *workshop* peran statistika data sains untuk kelas XII SMA Waskito lebih dari 70 secara signifikan.

BAHAN DAN METODE

Metode penelitian yang digunakan untuk keperluan analisa data adalah penelitian *Mixed-Method* (kualitatif-kuantitatif) sesuai (John, 2009). Analisa data kuantitatif menggunakan statistika inferensia. Jenis statistika inferensia yang digunakan untuk menjawab hipotesis adalah *One Mean t-Test* (Weiss, 2012; Sheskin, 2000). Sampel yang digunakan untuk analisa data, merupakan kelas XII SMA Waskito sebanyak 176 siswa-siswi. Adapun langkah atau prosedur untuk menjawab hipotesa sesuai dengan (Weiss, 2012).

Asumsi yang digunakan:

1. Sampel acak sederhana. Jelas, dalam PKM ini sampel yang mengikuti tes adalah *random*.
2. Sampel besar atau populasi berdistribusi normal. Karena sampel lebih dari 30, maka kategori sampel besar. Hal ini berdasar referensi (Mendenhall et all, 2013) yang mengatakan secara konservatif ukuran sampel besar adalah $n \geq 30$.
3. σ tidak diketahui (standar deviasi populasi tidak diketahui).

Step 1. Hipotesa.

$$H_0 : \mu = \mu_0$$

$$H_a : \mu > \mu_0$$

Step 2. Tulis α

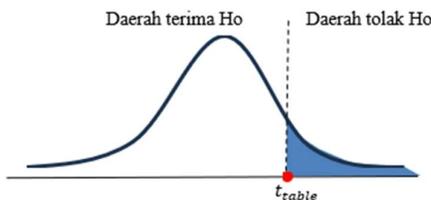
Step 3. Hitung nilai t_{stat}

$$t_{stat} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

\bar{x} merupakan rata-rata data sampel, μ_0 nilai yang diuji yaitu dalam PKM ini adalah $\mu_0 = 70$, s merupakan standar deviasi sampel dan n banyak data sampel.

Step 4. Hitung *Critical Value* (t_{tabel}).

Step 5. Bandingkan nilai t_{stat} dan t_{tabel} .



Gambar 2. Daerah penerimaan dan penolakan H_0

Jika nilai t_{stat} jatuh di area gelap (warna biru) maka tolak H_0 atau terima H_a . Sebaliknya, terima H_0 atau tolak H_a .

Step 6. Kesimpulan.

HASIL DAN PEMBAHASAN

Kegiatan *workshop* edukasi dilaksanakan sesuai permintaan dan jadwal yang telah diberikan oleh pihak sekolah Waskito. Tanggal pelaksanaan adalah 23 Februari 2024, dari mulai pukul 07.45 WIB hingga 10.30 WIB. *Workshop* berjalan dengan lancar, meskipun di awal terdapat satu kendala seperti perangkat proyektor dari sekolah yang tidak bisa terhubung dengan laptop. Namun, kendala ini tidak mengganggu proses PKM. Proyektor yang digunakan awalnya ada dua, namun karena satu terjadi kendala, yang digunakan adalah satu proyektor.

Selama pelaksanaan kegiatan pengabdian ini, tim menambah anggota untuk menjadi kameramen. Awal dari rencana adalah kameramen akan ditugaskan secara bergantian oleh Wiwik dan Ferdi, namun ketika melihat dan menimbang informasi dari pihak sekolah bahwa peserta yang hadir adalah seluruh kelas XII, maka rencana diubah dengan menambahkan anggota kameramen. Adapun anggota yang dimasukkan di hari pelaksanaan kegiatan *workshop* adalah mahasiswa prodi Statistika Universitas Matana, yaitu Calvin Riswandi dan Jeremy Heriyandi Saudi. Selanjutnya, untuk semua dokumentasi video dan foto selama kegiatan PKM dilaksanakan sudah terlampir pada *logbook* harian PKM yang telah disusun.

Agenda pengabdian dalam bentuk *workshop* edukasi ini dapat dikatakan sesuai dengan rencana, meskipun ada beberapa agenda yang menyesuaikan dengan keadaan asli di lapangan. Pada awal kegiatan, siswa-siswi kelas XII SMA Waskito memasuki Aula sekolah Waskito dengan mencatatkan tanda tangan peserta, seperti gambar 3.

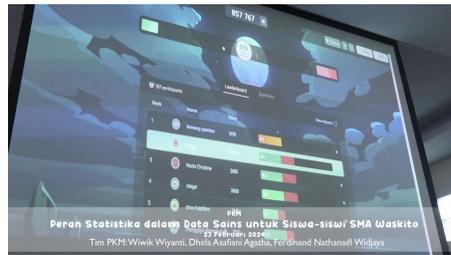


Gambar 3. Siswa melakukan tanda tangan kehadiran

Selanjutnya, acara dimulai dengan menyanyikan lagu Indonesia raya dan himne Waskito, gambar 4. Kemudian siswa mengerjakan soal *pre-test* dengan menggunakan Quiziz, seperti pada gambar 5, dilanjutkan oleh pemberian materi 1 oleh Wiwik Wiyanti, M.Sc mengenai *basic* ilmu peran statistika untuk data sains, seperti pada gambar 6.



Gambar 4. Seluruh peserta menyanyikan lagu Indonesia Raya dan himne Waskito



Gambar 5. Memanfaatkan teknologi (*Quiziz*)



Gambar 6. Materi dasar peran statistika untuk data sains

Setelah pemberian materi secara teori oleh pemateri 1, selanjutnya *ice breaking* dengan berbagai permainan untuk *refreshing* atau penyegaran para peserta *workshop*, tentu saja dilakukan sebelum belajar pengenalan *software* R untuk data sains. Adapun gambaran *ice breaking* adalah seperti pada gambar 7.



Gambar 7. *Ice breaking*

Acara selanjutnya diisi oleh pemateri kedua, yaitu Ferdinand Nathanael Wijaya dengan pengenalan *software* R dengan praktik untuk analisa data sains. Pada sesi ini, kendala terjadi. Proyektor yang aktif adalah 1 buah, yaitu sebelah kanan, sehingga peserta bagian kiri tidak dapat dengan jelas berpartisipasi mengikuti praktik. Hal ini dijadikan pertimbangan untuk PKM selanjutnya, supaya menggunakan laboratorium komputer apabila *workshop*

berhubungan dengan praktik dengan *software*. Adapun gambaran pemberian material R oleh pemateri 2, dapat dilihat pada gambar 8.



Gambar 8. Pemateri 2 mengenalkan *software* R

Selanjutnya adalah melakukan cek pemahaman peserta mengenai materi yang dipaparkan, salah satunya adalah dengan memberikan pertanyaan kuis dan pemberian hadiah. Gambaran nuansa pemberian pertanyaan dan hadiah adalah seperti pada gambar 9. Dilanjutkan dengan melaksanakan pemberian *post-test* dengan menggunakan Quiziz kepada seluruh peserta.



Gambar 9. Siswa antusias menjawab pertanyaan

Adapun dari rangkaian seluruh kegiatan pengabdian kali ini, dapat dikatakan sukses sesuai dengan rencana. Adapun kekurangan atau masalah kecil yang ditemukan ketika proses *workshop* berlangsung, tidak mengurangi kesuksesan acara.

Selanjutnya untuk analisa data secara kuantitatif. Pada awal proposal diusulkan dengan *paired sample t test*, akan tetapi berdasarkan data yang diperoleh, analisa ini tidak dimungkinkan untuk digunakan dengan alasan banyak ketidak sesuaian antara nama sampel yang mengikuti *pre test* maupun *post-test*. Beberapa siswa yang namanya ada di *pre-test*, di *post-test* tidak ada atau diduga berganti nama. Sehingga sulit untuk melacak manakah orang yang sama. Karena kendala ini, metode analisa data kuantitatif pada proposal diganti dengan menggunakan uji *t* satu kelompok sampel (*One-Mean t-Test*) dengan data nilai yang dianalisa adalah data hasil *post-test*.

Sebelum menggunakan metode (*One-Mean t-Test*), dilakukan terlebih dahulu *cleaning data*. *Cleaning data* adalah membuang data ganda ataupun data yang tidak mengikuti *post-test*. Adapun dari data *posttest* yang terkumpul, sebanyak 211 orang, terdapat

30 orang dengan data ganda, 5 data tidak mengikuti post-test, 1 siswa menjawab 1 soal, 3 siswa menjawab 2 soal, 2 siswa menjawab 3 soal, 4 siswa menjawab 4 soal, 1 siswa menjawab 7 soal, 2 siswa menjawab 8 soal, 3 siswa menjawab 9 soal dan 160 siswa menyelesaikan 10 soal. Adapun rangkuman hasilnya seperti tabel 1.

Tabel 1. Kelompok siswa berdasar banyak pertanyaan yang dijawab

Banyak soal yang dijawab	Frekuensi (siswa)
10	160
9	3
8	2
7	1
6	0
5	0
4	4
3	2
2	3
1	1
0	5
Data ganda	30
Total	211

Untuk keperluan analisa data, maka yang dianalisa semua siswa berbeda yang mengikuti *posttest*, yaitu sebanyak $211-30-5=176$ siswa. Adapun hipotesa yang diuji adalah apakah *workshop* yang diberikan secara kuantitatif secara signifikan dikatakan mempunyai nilai rata-rata melebihi 70 atau tidak, prosedur yang dilakukan sesuai dengan [7], yaitu:

Asumsi:

1. Sampel random. Jelas, dalam PKM ini sampel yang mengikuti tes adalah *random*
2. Sampel besar atau populasi normal. Karena sampel lebih dari 30, maka kategori sampel besar.
3. σ tidak diketahui (variansi data populasi dari PKM ini tidak diketahui)

Step 1. Hipotesa.

$$H_0 : \mu = \mu_0$$

$$H_a : \mu > \mu_0$$

Dalam penelitian ini $\mu_0 = 70$ yaitu mengacu nilai B atau Baik acuan penilaian di universitas Matana (nilai B adalah $70 \leq \text{nilai} < 80$) Sehingga, hipotesa menjadi

$$H_0 : \mu = 70$$

$$H_a : \mu > 70$$

Step 2. Tulis α (dalam analisa data PKM ini menggunakan $\alpha = 5\%$)

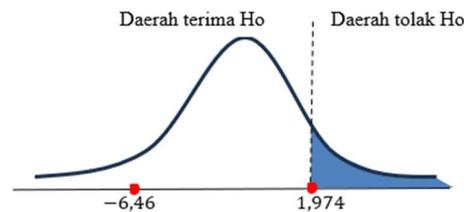
Step 3. Hitung nilai t_{stat}

$$t_{stat} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Dalam PKM ini, diperoleh bahwa $\bar{x} = 57,61$; $n = 176$; $\mu_0 = 70$; $s = 25,43$, sehingga $t = -6,46$

Step 4. Hitung *Critical Value* t_{tabel} . Nilai $t_{table} = t_{\alpha,df} = t_{5\%,(176-1)} = 1,974$

Step 5. Bandingkan nilai t_{test} dan t_{tabel} .



Karena nilai $t_{hit} = -6,46$ jatuh di area terang, maka terima H_0 .

Step 6. Kesimpulan.

Dari data *post-test* yang diperoleh tidak cukup bukti untuk mengatakan bahwa rata-rata nilai *post-test* lebih besar dari 70.

KESIMPULAN DAN SARAN

Hasil analisa secara kualitatif, secara keseluruhan kegiatan *workshop* berjalan sesuai dengan rencana meskipun ada beberapa kendala, namun tidak mengurangi kegiatan yang telah disusun pada rancangan awal. Secara kuantitatif, diperoleh kesimpulan bahwa nilai rata-rata pembelajaran melalui *workshop* peran statistika data sains untuk kelas XII SMA Waskito tidak lebih dari 70 secara signifikan.

Saran untuk PKM berikutnya adalah ketika berhubungan dengan praktikum, maka disarankan menggunakan laboratorium komputer supaya materi tepat guna untuk setiap individu siswa serta siswa dapat praktikum secara langsung. Selanjutnya adalah penggunaan Quizziz dapat meminta siswa untuk menggunakan nama asli, dikarenakan adanya ketidakkonsistenan nama saat *pre-test* dan *post-test*.

DAFTAR PUSTAKA

- Achmad, A. (2020). *Modul pembelajaran sma matematika umum*. Kementerian Pendidikan dan Kebudayaan Direktorat Jenderal Pendidikan Anak Usia Dini, Pendidikan Dasar dan Pendidikan Menengah Direktorat Sekolah Menengah Atas 2020. https://repositori.kemdikbud.go.id/21944/1/XII_Matematika-Umum_KD-3.2_Final.pdf (diakses 10 februari 2024)
- Ahdiat, A. (2023). Manfaat utama analisis data untuk bisnis menurut responden. Dapat diakses pada: <https://databoks.katadata.co.id/datapublish/2023/07/17/apa-manfaat-data-untuk-bisnis-ini-hasil-surveinya> (diakses 11 Januari 2024)
- Dirwan, N.A. (2024). <https://valiance.ai/data-scientist-adalah-pekerjaan-terseksi-di-abad-21/> (diakses 11 januari 2024)

- <https://dqlab.id/alasan-mengapa-profesi-data-scientist-indonesia-langka>. (2020). (Diakses 12 Januari 2024)
- John W, Creswell. (2009). *Research desain: qualitative, quantitative and mixed methods approach*. Third edition. New York: SAGE Publications, Inc;
- Mendenhall, W.III, Beaver, R.J, & Beaver, B.M. (2013). *Introduction to probability and statistics*. 14th Ed. Brooks/Cole:USA. ISBN-13: 978-1-133-10375-2
- Mursid, F. (2022). Kominfo:kebutuhan sdm praktisi data semakin meningkat pelatihan data science sejak usia dini, para pelajar akan memiliki kecakapan digital. Dapat diakses pada: <https://ekonomi.republika.co.id/berita/rfks7n380/kominfo-kebutuhan-sdm-praktisi-data-semakin-meningkat> (diakses 11 Januari 2024)
- Sheskin, D.J. (2000). *Handbook of Parametric and Nonparametric Statistical Procedures*. Second edition. United State of America: Chapman&Hall/CRC
- Unej, H. (2024). Indonesia butuh ribuan data scientist. dapat diakses pada: <https://unej.ac.id/blog/2019/12/15/indonesia-butuh-ribuan-data-scientist/> (diakses 12 Januari 2024)
- Weiss, N.A. (2012). *Introductory statistics*. Ninth edition. Pearson Education, Inc. ISBN 978-0-321-69122-4
- Wiyanti, W. (2017). Upaya meningkatkan kemampuan berfikir logis matematis mahasiswa papua dalam pemecahan masalah luas daerah di bawah kurva normal. *Jurnal derivat*. DOI: <https://doi.org/10.31316/j.derivat.v4i2.156>