

Jarot Budiasto
Dodi
Ketrin R. Manullang
Lawrence Adi Supriyono
Nengah Widya Utami

DATA ANALYTICS: MENJADI AHLI DALAM MENGOLAH DAN MENGANALISIS DATA



DATA ANALYTICS:

**Menjadi Ahli dalam Mengolah
dan Menganalisis Data**

Jarot Budiasto

Dodi

Ketrin R. Manullang

Lawrence Adi Supriyono

Nengah Widya Utami



GETPRESS INDONESIA

**DATA ANALYTICS: Menjadi Ahli dalam Mengolah
dan Menganalisis Data**

Penulis :

Jarot Budiasto

Dodi

Ketrin R. Manullang

Lawrence Adi Supriyono

Nengah Widya Utami

ISBN : 978-623-125-823-6

Editor : Mila Sari, M. Si.

Desain Sampul dan Tata Letak : Tri Putri Wahyuni, S. Pd.

PENERBIT : GET PRESS INDONESIA

Anggota IKAPI No. 033/SBA/2022

Jl. Palarik RT 01 RW 06 Kelurahan Air Pacah

Kecamatan Koto Tangah Padang Sumatera Barat

website: www.getpress.co.id

email: adm.getpress@gmail.com

Cetakan pertama, Juni 2025

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk
dan dengan cara apapun tanpa izin tertulis dari penerbit.

KATA PENGANTAR

Puji dan syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya sehingga buku berjudul *Data Analytics: Menjadi Ahli dalam Mengelolah dan Menganalisis Data* ini dapat diselesaikan. Buku ini hadir sebagai bentuk respons terhadap perkembangan teknologi dunia modern. Isi buku ini mencakup pembahasan tentang dasar-dasar data analytics, sumber dan pengolahan data, teknik analisis data, visualisasi data dan pengambilan keputusan, dan penerapan data analytics di dunia nyata.

Kami berharap buku ini dapat menjadi referensi yang bermanfaat dalam mendukung pengembangan Analytics Data di Indonesia. Akhir kata, semoga buku ini dapat memberikan manfaat dan inspirasi bagi semua pembaca.

Jakarta, Mei 2025

Penulis

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI.....	ii
DAFTAR GAMBAR	iv
DAFTAR TABLET	v
BAB 1 DASAR-DASAR <i>DATA ANALYTICS</i>.....	1
1.1 Pendahuluan: Mengapa <i>Data Analytics</i> Penting?	1
1.2 Pengenalan <i>Data Analytics</i>	5
1.3 Jenis-Jenis <i>Data Analytics</i>	11
1.4 <i>Tools</i> dan Teknologi Dasar	17
1.5 Studi Kasus: Kesuksesan <i>Data Analytics</i> di Dunia Nyata.....	23
1.6 Penutup: Langkah Awal Menjadi Ahli <i>Data Analytics</i> ..	26
DAFTAR PUSTAKA	31
BAB 2 SUMBER DAN PENGOLAHAN DATA.....	38
2.1 Pengantar Data.....	38
2.2 Sumber Data.....	41
2.3 Pengolahan Data.....	46
DAFTAR PUSTAKA	53
BAB 3 TEKNIK ANALISIS DATA	54
3.1 Pendahuluan	54
3.2 Definisi Teknik Analisis Data.....	55
3.3 Langkah-langkah Teknik Analisis Data	56
3.4 Jenis-Jenis Teknik Analisis Data	57
3.4 Perbedaan Teknik Analisis Data Kuantitatif dan Teknik Analisis Data Kualitatif.....	61
DAFTAR PUSTAKA	63
BAB 4 VISUALISASI DATA DAN PENGAMBILAN KEPUTUSAN	65
4.1 Pendahuluan	65
4.2 Tujuan Visualisasi Data.....	67
4.3 Prinsip-Prinsip Visualisasi Data yang Efektif.....	69
4.4 Alat-Alat untuk Visualisasi Data.....	71

4.5 Teknik Visualisasi Data	79
4.7 Manfaat Visualisasi Data dalam Pengambilan Keputusan	85
4.8 Tantangan dalam Visualisasi Data	89
4.10 Kesimpulan	92
DAFTAR PUSTAKA	94
BAB 5 PENERAPAN <i>DATA ANALYTICS</i> DI DUNIA NYATA	97
5.1 Pendahuluan	97
5.2 Penerapan <i>Data Analytics</i> pada Berbagai Sektor	98
5.3 Tantangan dalam Implementasi <i>Data Analytics</i>	116
DAFTAR PUSTAKA	117
BIODATA PENULIS	119

DAFTAR GAMBAR

Gambar 1. 1 Siklus <i>Data Analytics</i>	8
Gambar 1. 2 Progresi <i>Data Analytics</i>	16
Gambar 1. 3 Perbandingan Platform Cloud untuk <i>Data Analytics</i>	22
Gambar 4. 1 Contoh Grafik Garis.....	72
Gambar 4. 2 Contoh Grafik Batang.....	73
Gambar 4. 3 Contoh Grafik Diagram Lingkaran.....	73
Gambar 4. 4 Contoh Grafik Histogram.....	74
Gambar 4. 5 Contoh Grafik <i>Heatmap</i>	76
Gambar 4. 6 Contoh Dashboard Interaktif.....	77
Gambar 4. 7 Contoh Diagram Jaringan	78
Gambar 5. 1 Aplikasi Segmentasi Pelanggan	100
Gambar 5. 2 Aplikasi Analisis Sentimen.....	101
Gambar 5. 3 Aplikasi Analisis Tren Pasar.....	103
Gambar 5. 4 Aplikasi Diagnosis Penyakit	104
Gambar 5. 5 Aplikasi Prediksi Penyebaran Penyakit	106
Gambar 5. 6 Aplikasi <i>Fraud Detection</i>	107
Gambar 5. 7 Aplikasi Analisis Kredit	108
Gambar 5. 8 Aplikasi Tren Pasar Saham	110
Gambar 5. 9 Aplikasi Analisis Performa Siswa.....	111
Gambar 5. 10 Aplikasi Rekomendasi Kursus	112
Gambar 5. 11 Aplikasi Perencanaan Transportasi	113
Gambar 5. 12 Aplikasi Deteksi Anomali Anggaran Negara	114
Gambar 5. 13 Aplikasi Prediksi Kriminalitas	115

DAFTAR TABEL

Tabel 1. 1 Perbedaan *Data Analytics*, *Data Science*, dan
Business Intelligence 6

Tabel 1. 2 Perbandingan Jenis-Jenis *Data Analytics* 15

Tabel 1. 3. Perbandingan Platform Cloud untuk *Data*
Analytics 22

Tabel 1. 4 *Hard Skills* dan *Soft Skills* dalam *Data Analytics*..... 27

BAB 1

DASAR-DASAR *DATA ANALYTICS*

Oleh Jarot Budiasto

1.1 Pendahuluan: Mengapa *Data Analytics* Penting?

Dalam dunia modern yang dipenuhi oleh data dari berbagai sumber—media sosial, transaksi digital, sensor IoT, hingga catatan medis—kemampuan untuk memahami dan mengolah data menjadi hal yang tak lagi opsional, tetapi esensial. *Data analytics*, atau analitik data, kini telah menjadi bagian integral dari proses pengambilan keputusan di berbagai sektor. Kita sedang menyaksikan pergeseran besar dalam cara individu, organisasi, dan bahkan negara membuat keputusan: dari pendekatan berbasis intuisi atau pengalaman semata, menuju pendekatan berbasis bukti yang diperoleh dari data yang dianalisis secara sistematis.

1.1.1 Era *Data-Driven Decision Making*

Era *data-driven decision making* menandai sebuah revolusi dalam strategi organisasi. Dalam buku *Competing on Analytics* karya Davenport and Harris (2007), ditegaskan bahwa organisasi yang unggul dalam kompetisi global saat ini adalah mereka yang tidak hanya mengumpulkan data, tetapi juga mengintegrasikan analitik ke dalam seluruh proses bisnis mereka. Hal ini memungkinkan organisasi untuk

tidak hanya melihat apa yang terjadi, tetapi juga memahami mengapa itu terjadi, apa yang kemungkinan akan terjadi, dan bahkan apa yang sebaiknya dilakukan selanjutnya.

Organisasi seperti Netflix, Amazon, dan Google telah membuktikan kekuatan pendekatan berbasis data dalam mendorong inovasi dan efisiensi. Netflix, misalnya, tidak lagi mengandalkan intuisi untuk memilih konten yang diproduksi atau ditayangkan. Mereka menggunakan analitik prediktif berbasis data histori penonton, interaksi pengguna, hingga waktu tonton untuk menentukan jenis film atau serial yang kemungkinan besar akan sukses di pasar (Pham and Wan, 2023). Hasilnya, serial seperti *House of Cards* dan *Stranger Things* bukan hanya sukses besar, tetapi juga diciptakan berdasarkan hasil analitik, bukan sekadar kreativitas semata.

Amazon mengambil pendekatan yang serupa, tetapi dalam konteks perdagangan daring. Melalui data pelanggan yang dikumpulkan dari pencarian, pembelian, hingga ulasan produk, Amazon mampu memberikan rekomendasi produk yang dipersonalisasi, mengoptimalkan rantai pasok secara dinamis, dan memperkirakan permintaan pasar di masa mendatang (Kim, 2023). Sementara itu, Google mengandalkan *data analytics* untuk mengelola ratusan juta pencarian setiap hari, mengoptimalkan performa iklan berbasis preferensi pengguna, dan menyempurnakan hasil pencarian melalui pembelajaran mesin.

Pendekatan ini tidak hanya menguntungkan dari sisi ekonomi, tetapi juga menciptakan pengalaman yang lebih baik bagi pelanggan. Pada titik ini, *data analytics*

bukan sekadar alat bantu, melainkan menjadi fondasi strategi bisnis yang kompetitif dan berkelanjutan.

1.1.2 Perkembangan *Data Analytics* dalam Dekade Terakhir

Sepuluh tahun terakhir telah menjadi saksi dari transformasi besar dalam dunia analitik. Jika sebelumnya analisis data terbatas pada penyusunan laporan dan grafik statis menggunakan spreadsheet, kini kita telah memasuki era *advanced analytics* yang melibatkan *big data*, *machine learning*, dan kecerdasan buatan (AI).

Transformasi ini tidak terlepas dari peningkatan eksponensial volume data yang dihasilkan setiap hari. Mayer-Schönberger and Cukier (2013) dalam bukunya *Big Data: A Revolution That Will Transform How We Live, Work, and Think* menjelaskan bahwa pertumbuhan data tidak hanya dalam jumlah, tetapi juga dalam kecepatan, variasi, dan kompleksitas. Data tidak lagi terbatas pada angka, tetapi juga mencakup teks, gambar, suara, dan data sensor yang harus diproses dalam waktu nyaris seketika.

Pergeseran ini melahirkan kebutuhan akan teknologi dan pendekatan baru yang mampu menangani data dalam skala besar dan kompleks. Misalnya, teknologi *cloud computing* memungkinkan organisasi menyimpan dan mengolah data dalam jumlah besar tanpa harus memiliki infrastruktur fisik yang mahal. Di sisi lain, algoritma pembelajaran mesin dapat mengenali pola tersembunyi dalam data untuk tujuan prediksi dan klasifikasi, memberikan keunggulan kompetitif yang sebelumnya sulit dicapai.

Perubahan ini juga melahirkan berbagai disiplin ilmu baru seperti *data science*, *data engineering*, dan *AI ethics*, yang memperluas cakupan dan dampak *data analytics* di berbagai bidang, mulai dari layanan kesehatan yang lebih personal, prediksi bencana alam, hingga kebijakan publik berbasis data.

1.1.3 Tujuan dan Manfaat Mempelajari *Data Analytics*

Mempelajari *data analytics* bukan sekadar belajar mengolah angka, tetapi memahami cara mengubah data menjadi *insight* yang bermakna. Dengan menguasai keterampilan ini, seseorang mampu mengidentifikasi pola tersembunyi, menjelaskan hubungan sebab-akibat, memprediksi kemungkinan kejadian di masa depan, dan merekomendasikan tindakan yang paling optimal.

Bagi individu, penguasaan *data analytics* dapat membuka peluang karir yang sangat luas. Perusahaan di bidang teknologi, perbankan, kesehatan, pemerintahan, pendidikan, logistik, hingga hiburan, semuanya membutuhkan analisis data untuk membantu mereka tetap relevan dan kompetitif di era digital. Menurut laporan *The Future of Jobs Report 2025* yang dirilis oleh *World Economic Forum* (WEF), profesi seperti *data analyst* dan *data scientist* termasuk dalam jajaran pekerjaan dengan pertumbuhan tercepat dan permintaan tinggi di dekade ini (World Economic Forum, 2025).

Dari sisi organisasi, *data analytics* dapat membantu meningkatkan efisiensi operasional, menurunkan biaya, meningkatkan kepuasan pelanggan, serta mendorong inovasi produk dan layanan (Krishnamoorthy *et al.*, 2024; Saleh *et al.*, 2024; Joshi, Spilbergs and Mikelsone,

2025). Misalnya, perusahaan manufaktur menggunakan *data analytics* untuk memprediksi kerusakan mesin sebelum terjadi, sehingga dapat mengurangi waktu henti produksi dan menghemat biaya pemeliharaan (Kusumaningrum, Kurniati and Santosa, 2021).

Di dunia akademik dan pemerintahan, *data analytics* berperan penting dalam menyusun kebijakan berbasis bukti, mengelola data populasi, hingga mengoptimalkan anggaran pendidikan dan kesehatan publik. Dalam konteks yang lebih luas, kemampuan ini menjadi pilar penting dalam mewujudkan pemerintahan digital (*e-Government*), kota pintar (*smart city*), dan masyarakat berbasis pengetahuan (*knowledge-based society*).

Dengan kata lain, memahami dan menguasai *data analytics* adalah langkah awal untuk menjadi pengambil keputusan yang lebih bijaksana, lebih akurat, dan lebih adaptif dalam menghadapi kompleksitas dunia modern.

1.2 Pengenalan *Data Analytics*

Di tengah gempuran informasi digital yang semakin deras, kemampuan untuk membaca, memahami, dan memanfaatkan data menjadi sebuah kebutuhan yang tidak lagi bisa ditawar. Kehadiran *data analytics* menjembatani kebutuhan tersebut dengan menghadirkan pendekatan sistematis untuk mengubah data mentah menjadi wawasan yang bernilai. Dalam konteks inilah, *data analytics* menjadi landasan utama dalam pengambilan keputusan berbasis bukti (*evidence-based decision making*) di berbagai sektor, mulai dari bisnis, pendidikan, kesehatan, hingga pemerintahan.

1.2.1 Definisi dan Ruang Lingkup *Data Analytics*

Secara umum, *data analytics* dapat didefinisikan sebagai proses pengumpulan, pengolahan, dan analisis data dengan tujuan untuk menemukan pola, mengidentifikasi tren, dan menghasilkan *insight* yang mendukung pengambilan keputusan strategis (Igulu, Osuigbo and Singh, 2023). Proses ini melibatkan kombinasi antara teknik statistik, metode komputasi, dan pemahaman terhadap konteks data itu sendiri. Tidak hanya berfungsi sebagai alat analisis, *data analytics* juga berperan dalam mendesain strategi, meningkatkan efisiensi operasional, dan menciptakan inovasi berbasis data.

Provost and Fawcett (2013) dalam bukunya *Data Science for Business* menjelaskan bahwa *data analytics* merupakan bagian dari ekosistem ilmu data (*data science*), namun memiliki fokus yang lebih aplikatif dalam mengolah dan menyajikan data sebagai dasar pengambilan keputusan. Meskipun sering disandingkan dengan *data science* dan *business intelligence*, ketiganya memiliki perbedaan mendasar, seperti dirangkum dalam Tabel 1.1.

Tabel 1. 1 Perbedaan *Data Analytics*, *Data Science*, dan *Business Intelligence*

Fokus Utama



Data Analytics

Menganalisis data untuk menemukan pola dan mendukung pengambilan keputusan. Bersifat aplikatif dan berorientasi pada solusi praktis untuk masalah nyata.

Fokus Utama



Data Science

Disiplin ilmiah yang mencakup proses pengumpulan, analisis, interpretasi, hingga prediksi data. Cakupannya lebih luas dan mendalam, termasuk riset dan pengembangan model algoritmik.



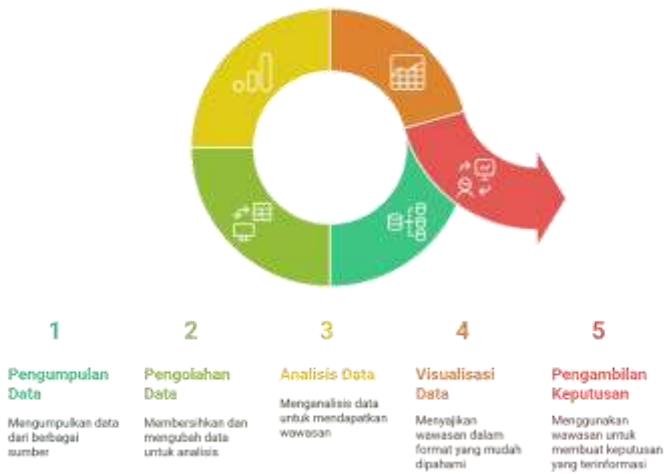
Business
Intelligence (BI)

Menyajikan data historis dalam bentuk laporan dan dashboard untuk kebutuhan bisnis. Fokus utamanya adalah pelaporan dan visualisasi data masa lalu untuk mendukung operasi bisnis sehari-hari.

Dengan demikian, *data analytics* menempati posisi strategis sebagai jembatan antara data dan keputusan, yang dapat beroperasi baik dalam kerangka kerja BI maupun menjadi bagian dari proses *data science* yang lebih kompleks.

1.2.2 Komponen Utama *Data Analytics*

Untuk memahami bagaimana *data analytics* bekerja secara menyeluruh, penting untuk mengenal komponen-komponen utama yang membentuk keseluruhan proses analitik. Komponen-komponen ini tidak berdiri sendiri, melainkan saling berinteraksi dalam siklus analisis yang berulang dan berkelanjutan.



Gambar 1. 1 Siklus *Data Analytics*

1. Pengumpulan Data (*Data Collection*)

Tahap awal dalam proses *data analytics* dimulai dengan pengumpulan data dari berbagai sumber. Sumber data bisa bersifat internal (misalnya: sistem informasi perusahaan, log transaksi penjualan) maupun eksternal (misalnya: media sosial, sensor IoT, data publik) (Jurnaidi *et al.*, 2025). Keberhasilan analisis sangat bergantung pada kualitas dan relevansi data yang dikumpulkan. Oleh karena itu, proses ini sering melibatkan teknik validasi awal untuk memastikan integritas dan konsistensi data.

2. Pengolahan Data (*Data Processing*)

Setelah data berhasil dikumpulkan, langkah berikutnya adalah membersihkannya dari ketidakteraturan seperti duplikasi, nilai hilang, atau kesalahan input (Dhaouadi *et al.*, 2019). Proses ini

disebut *data cleaning*, yang merupakan salah satu tahapan paling krusial dalam analitik data. Selanjutnya, data dapat ditransformasikan dalam format yang lebih siap dianalisis, misalnya melalui normalisasi, penggabungan antar tabel, atau pengkodean variabel kategorikal.

3. Analisis Data (*Data Analysis*)

Inilah jantung dari aktivitas *data analytics*. Di tahap ini, data yang telah diproses dianalisis menggunakan berbagai teknik, tergantung pada tujuan yang ingin dicapai. Untuk menjawab pertanyaan "apa yang sedang terjadi?", analisis mungkin menggunakan teknik statistik deskriptif. Sementara untuk mengetahui "apa yang kemungkinan besar akan terjadi di masa depan?", digunakan pendekatan prediktif seperti regresi atau algoritma *machine learning* (Igulu, Osuigbo and Singh, 2023).

4. Visualisasi Data (*Data Visualization*)

Hasil dari analisis perlu disampaikan dalam bentuk yang mudah dipahami. Visualisasi data memungkinkan kita melihat pola dan *insight* dengan lebih cepat melalui grafik, peta, dashboard interaktif, atau diagram (Singh, Singh and Prabha, 2022; Budiasto, Tallulembang, *et al.*, 2025). *Tools* seperti Tableau, Power BI, dan Matplotlib banyak digunakan untuk keperluan ini karena mampu menyajikan data secara informatif dan menarik secara visual.

5. Pengambilan Keputusan (*Decision-Making*)

Akhir dari proses *data analytics* adalah penerapan hasil analisis sebagai dasar pengambilan keputusan. Dengan adanya data yang akurat dan analisis yang tepat, organisasi atau individu dapat membuat keputusan yang lebih terinformasi, minim risiko, dan berbasis pada realitas yang terukur, bukan sekadar intuisi.

1.2.3 Contoh Aplikasi *Data Analytics* dalam Kehidupan Sehari-hari

Meskipun terdengar kompleks, aplikasi *data analytics* sebenarnya sudah menjadi bagian dari kehidupan kita sehari-hari. Beberapa contoh berikut menggambarkan bagaimana data digunakan secara cerdas untuk meningkatkan kualitas layanan dan pengalaman pengguna:

- **Rekomendasi Produk di Platform E-Commerce**

Ketika pengguna mengunjungi situs seperti Tokopedia atau Shopee, sistem secara otomatis menampilkan produk-produk yang paling relevan dengan minat mereka. Hal ini bukan kebetulan, melainkan hasil dari algoritma *recommendation engine* yang menganalisis pola pembelian, riwayat pencarian, dan interaksi pengguna lain yang serupa (Sharma *et al.*, 2023).

- **Prediksi Cuaca dan Analisis Risiko Bencana**

Badan Meteorologi dan lembaga penanggulangan bencana menggunakan data dari satelit, sensor suhu dan kelembapan, serta catatan historis cuaca

untuk memprediksi terjadinya hujan deras, badai, atau potensi banjir (Choudha *et al.*, 2024). Analisis ini menjadi dasar dalam memberikan peringatan dini kepada masyarakat.

- **Personalisasi Konten di Media Sosial**

Platform seperti Instagram, TikTok, dan YouTube menggunakan *data analytics* untuk mempelajari perilaku pengguna: video yang sering ditonton, waktu menonton, atau jenis konten yang diabaikan. Dengan analisis ini, algoritma dapat menyajikan konten yang diprediksi paling sesuai dengan preferensi setiap individu, meningkatkan keterlibatan pengguna.

Contoh-contoh tersebut menunjukkan bahwa *data analytics* tidak hanya penting bagi perusahaan teknologi besar, tetapi juga berkontribusi langsung dalam memudahkan kehidupan masyarakat dan mendukung layanan publik yang lebih responsif dan adaptif.

1.3 Jenis-Jenis *Data Analytics*

Dalam era digital saat ini, keberadaan data yang melimpah menjadikan analisis data (*data analytics*) sebagai salah satu fondasi utama dalam pengambilan keputusan. Namun, tidak semua analisis data memiliki tujuan yang sama. Secara umum, terdapat empat jenis utama dari *data analytics*, yaitu *Descriptive Analytics*, *Diagnostic Analytics*, *Predictive Analytics*, dan *Prescriptive Analytics*. Masing-masing jenis ini memiliki fungsi dan pendekatan yang berbeda dalam memberikan wawasan dari data yang tersedia.

1.3.1 *Descriptive Analytics*: Memahami Apa yang Terjadi

Descriptive analytics merupakan jenis analisis data yang paling dasar, namun juga paling umum digunakan. Tujuan utama dari pendekatan ini adalah untuk memberikan gambaran tentang kejadian masa lalu berdasarkan data historis yang telah dikumpulkan (Geetha and Sujatha, 2024). Analisis deskriptif menjawab pertanyaan seperti "*Apa yang telah terjadi?*", dan berperan penting dalam menyajikan fakta secara ringkas dan jelas.

Contoh paling nyata dari *descriptive analytics* dapat ditemukan dalam laporan bulanan penjualan, dashboard kinerja perusahaan (*Key Performance Indicators*), statistik pengunjung website, serta data demografis pelanggan. Melalui penyajian data dalam bentuk tabel, grafik batang, diagram garis, atau *pie chart*, pengguna dapat dengan mudah mengidentifikasi tren dan pola yang muncul dari waktu ke waktu. Informasi ini sangat berharga karena menjadi dasar awal untuk analisis lanjutan.

1.3.2 *Diagnostic Analytics*: Mengetahui Mengapa Terjadi

Setelah mengetahui apa yang telah terjadi melalui analisis deskriptif, langkah selanjutnya adalah memahami mengapa hal tersebut terjadi. *Diagnostic analytics* hadir untuk mengisi kebutuhan ini. Pendekatan ini berfokus pada eksplorasi hubungan antar variabel serta identifikasi akar penyebab dari suatu fenomena.

Diagnostic analytics menjawab pertanyaan seperti "*Mengapa penjualan menurun?*" atau "*Mengapa tingkat kepuasan pelanggan menurun?*". Untuk menjawab pertanyaan tersebut, teknik seperti *drill-down analysis*,

data discovery, dan *correlation analysis* digunakan secara intensif. Misalnya, jika sebuah perusahaan ritel mengalami penurunan penjualan, *diagnostic analytics* dapat mengungkapkan bahwa penyebab utamanya adalah berkurangnya stok produk unggulan, promosi yang tidak berjalan optimal, atau munculnya pesaing baru di pasar.

Jenis analisis ini sangat berguna bagi manajer atau pengambil keputusan karena memberikan konteks yang lebih mendalam atas data historis, serta membantu menghindari pengambilan keputusan berdasarkan asumsi semata. Tanpa pemahaman yang benar atas penyebab masalah, tindakan korektif yang dilakukan bisa menjadi tidak tepat sasaran.

1.3.3 *Predictive Analytics*: Memprediksi Apa yang Akan Terjadi

Predictive analytics merupakan bentuk analisis yang lebih maju karena bertujuan untuk meramalkan apa yang akan terjadi di masa depan berdasarkan pola dan tren masa lalu (Kumar and L., 2018). Jenis analisis ini menggunakan teknik statistik, *machine learning*, dan *data mining* untuk membangun model prediktif. Model ini kemudian digunakan untuk memproyeksikan kejadian yang belum terjadi, dengan mempertimbangkan berbagai variabel yang relevan.

Sebagai contoh, perusahaan *e-commerce* menggunakan *predictive analytics* untuk memprediksi produk apa yang akan dibeli pelanggan berikutnya berdasarkan riwayat pembelian dan perilaku browsing mereka. Di sektor keuangan, *predictive analytics* digunakan untuk menilai risiko gagal bayar dari calon

peminjam (Wu, 2022). Dalam dunia kesehatan, model prediktif dapat digunakan untuk memprediksi kemungkinan kambuhnya suatu penyakit pada pasien tertentu (Wang *et al.*, 2020).

Menurut Siegel (2013) dalam bukunya *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*, kemampuan untuk memprediksi perilaku atau hasil secara akurat dapat menjadi kekuatan kompetitif yang sangat besar dalam bisnis modern. Namun, penting untuk diingat bahwa hasil dari *predictive analytics* bukanlah kepastian, melainkan estimasi probabilistik yang tetap memerlukan validasi dan pengawasan terus-menerus.

1.3.4 Prescriptive Analytics: Menentukan Apa yang Harus Dilakukan

Prescriptive analytics merupakan tahap paling lanjut dalam hierarki *data analytics*. Tujuan utamanya bukan hanya memahami masa lalu atau meramalkan masa depan, tetapi juga merekomendasikan tindakan terbaik yang seharusnya diambil. Analisis ini menjawab pertanyaan seperti "*Apa yang harus dilakukan selanjutnya?*" atau "*Keputusan mana yang akan menghasilkan hasil terbaik?*"

Untuk mencapai tujuannya, *prescriptive analytics* menggabungkan hasil prediksi dengan teknik optimasi, simulasi skenario, dan kadang-kadang *reinforcement learning* (Lepenioti *et al.*, 2020; Alkan, Menguc and Kabak, 2022). Pendekatan ini sangat berguna dalam konteks pengambilan keputusan yang kompleks, di mana terdapat berbagai opsi tindakan dengan dampak yang berbeda-beda.

Contoh implementasi *prescriptive analytics* adalah sistem rekomendasi di platform digital seperti Netflix atau Spotify, yang tidak hanya memprediksi apa yang mungkin disukai pengguna, tetapi juga menentukan konten apa yang sebaiknya ditampilkan terlebih dahulu. Di bidang logistik, *prescriptive analytics* digunakan untuk mengoptimalkan jalur pengiriman agar lebih efisien dari segi waktu dan biaya.

Jenis analisis ini sangat berguna dalam pengambilan keputusan strategis dan taktis karena memberikan panduan berbasis data yang dapat langsung diterapkan.

Tabel 1. 2 Perbandingan Jenis-Jenis *Data Analytics*

Jenis <i>Analytics</i>	Fokus Analisis	Contoh Aplikasi
<i>Descriptive</i>	Menyajikan data masa lalu. <i>Apa yang terjadi?</i>	Laporan penjualan, tren pengunjung website
<i>Diagnostic</i>	Menjelaskan penyebab. <i>Mengapa terjadi?</i>	Penyebab <i>churn</i> pelanggan, penurunan performa
<i>Predictive</i>	Meramalkan masa depan. <i>Apa yang akan terjadi?</i>	Prediksi permintaan, deteksi fraud
<i>Prescriptive</i>	Merekomendasikan tindakan. <i>Apa yang sebaiknya dilakukan?</i>	Penjadwalan produksi, strategi harga dinamis

Secara keseluruhan, keempat jenis *data analytics* ini dapat diibaratkan sebagai tahapan bertingkat dalam pengolahan dan pemanfaatan data. *Descriptive analytics* memberi pemahaman dasar tentang apa yang terjadi. *Diagnostic analytics* menjelaskan penyebab di balik kejadian tersebut. *Predictive analytics* membuka wawasan tentang kemungkinan yang akan datang. Dan *prescriptive analytics* membantu kita menentukan langkah paling bijak untuk menghadapi masa depan.

Setiap jenis memiliki peran dan kegunaannya masing-masing. Dalam praktik terbaik, organisasi yang cerdas akan mengintegrasikan semua jenis analisis ini untuk menciptakan sistem pengambilan keputusan yang komprehensif, adaptif, dan berkelanjutan.



Gambar 1. 2 Progresi *Data Analytics*

1.4 *Tools* dan Teknologi Dasar

Dalam dunia *data analytics*, pemahaman terhadap *tools* dan teknologi dasar bukanlah sebuah opsi, melainkan kebutuhan mendasar. *Tools* ini menjadi jembatan antara data mentah yang kompleks dengan *insight* bernilai yang mampu mendorong pengambilan keputusan yang lebih cerdas. Kemampuan untuk memilih dan mengoperasikan *tools* dengan tepat tidak hanya meningkatkan efisiensi analisis, tetapi juga menjadi indikator profesionalisme seorang analis data.

1.4.1 Pemilihan *Tools* Berdasarkan Kebutuhan

Pemilihan *tools* dalam *data analytics* tidak dapat dilakukan secara sembarangan. Setiap proyek analitik membawa tantangan unik, mulai dari volume data yang besar, variasi bentuk data, hingga kebutuhan analisis yang kompleks. Oleh karena itu, pemilihan alat perlu disesuaikan dengan beberapa faktor, seperti jenis data yang akan dianalisis (terstruktur atau tidak terstruktur), tingkat kompleksitas analisis (deskriptif, diagnostik, prediktif, atau preskriptif), skalabilitas, dan tujuan akhir dari analisis tersebut.

Sebagai contoh, ketika seorang analis bekerja dengan data terstruktur dalam jumlah kecil untuk keperluan pelaporan sederhana, aplikasi seperti Microsoft Excel sudah cukup memadai. Namun, untuk analisis prediktif berbasis *machine learning* terhadap jutaan data transaksi, dibutuhkan *tools* seperti Python yang mampu mengolah data besar dengan efisiensi tinggi. Dalam konteks ini, kekuatan Python terletak pada kemampuannya memproses data secara fleksibel dan

mendalam, melalui berbagai pustaka yang tersedia secara *open-source* (Budiasto, Ismanto, *et al.*, 2025).

Tidak kalah penting adalah mempertimbangkan keterampilan tim pengguna dan ketersediaan infrastruktur teknologi informasi. *Tools* yang terlalu kompleks bagi pengguna dapat memperlambat proses, bahkan menciptakan hambatan kolaborasi.

1.4.2 Pengenalan *Tools* Populer

Seiring berkembangnya teknologi, ekosistem *tools* untuk *data analytics* semakin luas dan beragam. Beberapa di antaranya telah menjadi standar industri dan banyak digunakan di berbagai sektor.

➤ Python dan R

Python dan R merupakan dua bahasa pemrograman utama yang mendominasi dunia *data analytics*. Python dikenal luas karena sintaksnya yang mudah dipelajari serta ekosistem pustaka yang sangat kaya, seperti Pandas untuk manipulasi data, Matplotlib dan Seaborn untuk visualisasi, serta Scikit-learn dan TensorFlow untuk *machine learning* (Budiasto, Tallulembang, *et al.*, 2025). Keunggulan Python juga terletak pada fleksibilitasnya yang memungkinkan integrasi dengan sistem lain seperti web, cloud, dan database.

Sementara itu, R lebih banyak digunakan di kalangan akademisi dan praktisi statistik. Bahasa ini memiliki kekuatan dalam visualisasi data statistik dan eksplorasi data, terutama melalui pustaka seperti ggplot2, dplyr, dan tidyr. Analisis

regresi, uji hipotesis, dan eksplorasi variabel sangat mudah dilakukan dalam R.

➤ **SQL (*Structured Query Language*)**

SQL tetap menjadi alat utama dalam dunia database. Bahasa ini digunakan untuk mengekstrak, memanipulasi, dan mengelola data dalam sistem manajemen basis data relasional (RDBMS) seperti MySQL, PostgreSQL, dan Microsoft SQL Server. Hampir semua data dalam organisasi besar tersimpan dalam basis data, menjadikan SQL sebagai keterampilan yang wajib dimiliki oleh analis data di berbagai level.

SQL sangat efektif untuk mengakses subset data secara cepat melalui teknik *filtering*, *grouping*, dan *aggregation*. Dengan menguasai SQL, seorang analis dapat melakukan eksplorasi awal data sebelum melanjutkan ke tahap analisis yang lebih mendalam.

➤ **Microsoft Excel dan Google Sheets**

Meski dianggap sebagai *tools* dasar, Excel dan Google Sheets tetap relevan dalam praktik *data analytics*, terutama untuk analisis sederhana, perhitungan statistik dasar, serta pembuatan grafik dan dashboard. Kelebihan utama Excel terletak pada antarmuka yang intuitif dan kemudahan dalam mengolah data kecil secara cepat. Excel juga mendukung penggunaan formula kompleks dan automasi sederhana melalui VBA (*Visual Basic for Applications*).

Google Sheets memberikan keunggulan dalam kolaborasi secara *real-time*, sehingga sangat cocok digunakan dalam tim yang bekerja secara terdistribusi.

➤ **Tableau dan Power BI**

Untuk kebutuhan visualisasi data yang menarik dan interaktif, Tableau dan Power BI menjadi dua *tools* yang sangat populer. Tableau dikenal karena kemampuannya membuat visualisasi data yang kompleks secara cepat tanpa perlu menulis kode. Power BI, yang dikembangkan oleh Microsoft, memiliki keunggulan integrasi dengan Microsoft Excel dan Azure, serta lebih ekonomis dalam implementasi di perusahaan skala menengah.

Visualisasi data memiliki peran penting dalam membantu pengambil keputusan memahami *insight* secara cepat dan intuitif. Oleh karena itu, keahlian menggunakan *tools* visualisasi menjadi nilai tambah yang besar dalam profil seorang analis data.

1.4.3 Lingkungan Pengembangan dan *Platform Cloud*

Dalam praktiknya, *data analytics* modern tidak terlepas dari penggunaan lingkungan pengembangan interaktif dan platform komputasi awan (*cloud*). Hal ini dikarenakan kebutuhan akan skalabilitas, kecepatan pemrosesan, serta kolaborasi dalam tim yang semakin kompleks.

➤ ***Jupyter Notebook***

Jupyter Notebook adalah lingkungan kerja berbasis web yang mendukung eksekusi kode Python secara interaktif, lengkap dengan visualisasi dan dokumentasi dalam satu file. Tool ini sangat populer di kalangan data scientist dan akademisi karena memungkinkan eksplorasi data, pengembangan model, dan pelaporan dalam satu tempat. Kode dapat dijalankan secara sel (*cell*), sehingga memudahkan debugging dan pengembangan bertahap.

➤ ***Google Colaboratory (Colab)***

Google Colab adalah versi cloud dari Jupyter Notebook yang dapat diakses secara gratis melalui akun Google. Salah satu keunggulannya adalah tersedianya GPU dan TPU untuk komputasi berbasis *machine learning* tanpa perlu instalasi lokal. Colab sangat cocok digunakan untuk pembelajaran, riset, maupun prototyping model skala kecil hingga menengah.

➤ ***Platform Cloud: AWS, GCP, dan Azure***

Untuk kebutuhan analisis data berskala besar, banyak organisasi kini beralih ke *platform cloud* seperti *Amazon Web Services (AWS)*, *Google Cloud Platform (GCP)*, dan *Microsoft Azure*. Layanan seperti AWS SageMaker, Google BigQuery, dan Azure Synapse Analytics memungkinkan analisis data yang kompleks dan deployment model *machine learning* secara efisien, cepat, dan aman.

Tabel 1. 3. Perbandingan Platform Cloud untuk *Data Analytics*

Fitur Utama	
 Amazon Web Services	Amazon Redshift, AWS Glue, AWS Athena, SageMaker Layanan sangat luas, dokumentasi lengkap, pasar terbesar Komunitas besar, ekosistem berkembang pesat
 Google Cloud Platform	BigQuery, Dataflow, Dataproc, Vertex AI Integrasi kuat dengan AI dan ML Google, performa BigQuery sangat cepat Dokumentasi kuat terutama untuk BigQuery dan AI
 Microsoft Azure	Azure Synapse Analytics, Data Factory, ML Studio Terintegrasi penuh dengan produk Microsoft seperti Excel dan Power BI Dukungan ekosistem Microsoft yang luas

Penggunaan cloud juga memberikan fleksibilitas dalam pengelolaan infrastruktur, terutama dalam hal penyimpanan data besar (*big data*), manajemen pipeline ETL (*Extract, Transform, Load*), dan automasi model analitik (Gupta, Mehrotra and Khan, 2015; Bhatlawande, Rajandekar and Shilaskar, 2024).

Secara keseluruhan, penguasaan terhadap berbagai *tools* dan teknologi dalam *data analytics* adalah langkah penting dalam membangun keahlian yang solid. *Tools* bukanlah tujuan akhir, melainkan sarana untuk memahami data dan menghasilkan *insight* yang bermakna. Oleh karena itu, penting bagi setiap calon

analisis untuk terus mengeksplorasi, belajar, dan menyesuaikan penggunaan *tools* sesuai kebutuhan dan perkembangan teknologi.

1.5 Studi Kasus: Kesuksesan *Data Analytics* di Dunia Nyata

Perkembangan pesat teknologi informasi dan ketersediaan data dalam jumlah besar telah mendorong berbagai organisasi untuk menerapkan *data analytics* sebagai bagian dari strategi inti mereka. Data tidak lagi hanya dianggap sebagai hasil sampingan dari proses bisnis, tetapi telah menjadi aset strategis yang menentukan arah kebijakan, efisiensi operasional, hingga kepuasan pelanggan. Beberapa perusahaan global terkemuka telah menunjukkan bagaimana *data analytics* dapat memberikan nilai nyata yang terukur. Berikut adalah tiga studi kasus dari perusahaan-perusahaan yang berhasil memanfaatkan data secara optimal untuk meraih keunggulan kompetitif.

1.5.1 Netflix: Personalisasi Konten dan Peningkatan Keterlibatan Pengguna

Netflix merupakan contoh klasik dari perusahaan yang berhasil mengintegrasikan *data analytics* ke dalam hampir seluruh aspek layanannya, terutama dalam hal personalisasi konten. Dengan lebih dari 230 juta pelanggan di seluruh dunia, Netflix mengumpulkan dan menganalisis berbagai jenis data perilaku pengguna, seperti riwayat tontonan, genre favorit, durasi menonton, serta interaksi terhadap suatu tayangan (seperti jeda, *rewind*, dan *skip*).

Data ini kemudian diolah menggunakan algoritma *machine learning* dan sistem rekomendasi berbasis kolaboratif (*collaborative filtering*) serta konten (*content-based filtering*). Hasilnya adalah pengalaman pengguna yang sangat personal—konten yang ditampilkan pada halaman beranda setiap pengguna disesuaikan secara khusus dengan preferensi individu mereka. Studi oleh Gomez-Uribe and Hunt (2016) menunjukkan bahwa lebih dari 80% konten yang ditonton di Netflix berasal dari sistem rekomendasi tersebut. Strategi ini terbukti mampu meningkatkan keterlibatan pengguna, memperpanjang durasi berlangganan, dan mengurangi tingkat churn.

Selain itu, Netflix juga menggunakan *data analytics* untuk memandu keputusan produksi konten orisinal. Salah satu contoh fenomenal adalah serial "House of Cards" yang dikembangkan berdasarkan analisis data pengguna tentang aktor, genre, dan sutradara yang paling disukai. Pendekatan ini memperkuat posisi Netflix sebagai pemimpin inovatif dalam industri hiburan digital.

1.5.2 Amazon: Efisiensi Operasional dan Peningkatan Pengalaman Pelanggan

Amazon dikenal sebagai perusahaan ritel daring terbesar di dunia yang sukses memanfaatkan *data analytics* dalam berbagai dimensi operasionalnya. Salah satu aspek yang paling menonjol adalah penerapan analitik prediktif dalam sistem rantai pasok dan logistik. Melalui pendekatan yang disebut *anticipatory shipping*, Amazon mampu memprediksi produk apa yang kemungkinan besar akan dibeli oleh pelanggan, dan mengirimkan produk tersebut ke gudang distribusi

terdekat bahkan sebelum transaksi dilakukan. Teknologi ini memungkinkan pengiriman lebih cepat dan kepuasan pelanggan yang lebih tinggi (Marr, 2016).

Menurut laporan McKinsey & Company dalam artikel *How retailers can keep up with consumers*, Amazon telah menerapkan sistem rekomendasi produk yang canggih. Sistem ini menggabungkan data historis pembelian pelanggan, ulasan produk, perilaku pencarian, dan pola pembelian pelanggan lain untuk menyarankan produk yang relevan. Efektivitas sistem ini terbukti dengan kontribusinya terhadap sekitar 35% dari total pendapatan Amazon (Mackenzie, Meyer and Noble, 2013).

Penggunaan *data analytics* yang mendalam telah menjadikan Amazon sebagai contoh unggulan dalam transformasi digital, membuktikan bahwa kecepatan dan akurasi dalam memahami perilaku pelanggan merupakan faktor kunci dalam memenangkan pasar digital yang kompetitif.

1.5.3 Uber: Pemanfaatan Data *Real-Time* untuk Efisiensi Layanan Transportasi

Uber, sebagai salah satu pelopor layanan transportasi berbasis aplikasi, mengandalkan *data analytics* dalam mengelola dan mengoptimalkan seluruh ekosistem layanannya. Aplikasi Uber secara *real-time* mengumpulkan data lokasi pengguna, kondisi lalu lintas, tingkat permintaan, dan waktu tempuh untuk mengatur ketersediaan kendaraan secara efisien (Fu and Soman, 2021). Salah satu inovasi utama yang lahir dari proses ini adalah sistem *surge pricing*—yaitu penyesuaian

harga dinamis berdasarkan fluktuasi permintaan dan ketersediaan pengemudi di suatu wilayah.

Melalui algoritma kompleks yang memproses data dalam hitungan detik, Uber mampu menjaga keseimbangan antara penumpang dan pengemudi. Ketika terjadi lonjakan permintaan di suatu area, harga secara otomatis meningkat untuk menarik lebih banyak pengemudi ke lokasi tersebut, sekaligus mengatur permintaan agar tidak melebihi kapasitas layanan. Penelitian oleh Chen, Mislove and Wilson (2015) menunjukkan bahwa strategi ini secara signifikan meningkatkan efisiensi dan kepuasan pengguna, meskipun sempat menuai kontroversi terkait keadilan harga.

Selain *pricing*, Uber juga memanfaatkan *data analytics* untuk rute perjalanan tercepat, prediksi waktu tempuh, dan bahkan evaluasi performa pengemudi dan pengguna. Dengan memanfaatkan miliaran titik data yang terus-menerus dihasilkan dari setiap perjalanan, Uber terus mengasah algoritmanya untuk memberikan layanan yang lebih cepat, nyaman, dan aman.

1.6 Penutup: Langkah Awal Menjadi Ahli *Data Analytics*

Menjadi seorang ahli dalam bidang *data analytics* bukanlah tujuan yang bisa dicapai dalam satu malam, melainkan perjalanan panjang yang membutuhkan kombinasi antara pengetahuan teoritis, keterampilan teknis, kemampuan berpikir kritis, serta pemahaman mendalam terhadap konteks bisnis dan sosial di mana data digunakan. Bab ini menutup pembahasan awal tentang fondasi dasar

data analytics, dan menjadi pintu gerbang menuju eksplorasi yang lebih luas dan mendalam.

Di era saat ini, data telah menjadi aset yang lebih berharga daripada minyak (Kropf, 2024). Perusahaan-perusahaan global dan institusi pemerintahan berlomba-lomba membangun sistem analitik yang canggih untuk memahami pasar, merespons perubahan, dan memprediksi masa depan. Dalam konteks inilah, peran seorang *data analyst*, *data scientist*, maupun *data engineer* menjadi sangat strategis. Namun, bagaimana sebenarnya langkah awal menjadi ahli di bidang ini? Apa saja yang harus dipelajari dan dikembangkan?

1.6.1 Keterampilan yang Dibutuhkan

Langkah awal menjadi ahli *data analytics* adalah memahami keterampilan apa saja yang perlu dikuasai. Secara umum, keterampilan ini terbagi menjadi dua: keterampilan teknis (*hard skills*) dan keterampilan non-teknis (*soft skills*), yang keduanya saling melengkapi (Zhang, Le and Chen, 2023; Levendis and Indika, 2025).

Tabel 1. 4 *Hard Skills* dan *Soft Skills* dalam *Data Analytics*

	Keterampilan	Deskripsi
Hard Skills	Statistik dan Matematika Dasar	Memahami distribusi, regresi, korelasi, dan probabilitas.
	Pemrograman (Python, R)	Menulis kode untuk analisis, pembersihan data, dan pemodelan.

	Keterampilan	Deskripsi
Soft Skills	Manipulasi dan Visualisasi Data	Mengolah data dan menyajikannya dalam bentuk grafis informatif.
	SQL (<i>Structured Query Language</i>)	Mengakses dan mengelola data dari basis data relasional.
	Berpikir Kritis dan <i>Problem Solving</i>	Menganalisis masalah dan mencari solusi berbasis data.
	Komunikasi Efektif	Menyampaikan temuan analitis secara jelas ke berbagai audiens.
	Kolaborasi Tim	Bekerja sama dalam tim lintas fungsi dan disiplin.
	Adaptasi terhadap Teknologi Baru	Cepat mempelajari <i>tools</i> dan teknologi baru dalam <i>data analytics</i> .

1.6.2 Sumber Belajar dan Komunitas

Seiring dengan meningkatnya minat terhadap bidang ini, tersedia banyak sumber belajar yang dapat diakses oleh siapa saja, kapan saja, di mana saja. Kemajuan teknologi telah membuka peluang belajar yang luas melalui berbagai platform daring.

Kursus daring seperti yang ditawarkan oleh Coursera, edX, dan Udemy menyediakan materi pembelajaran dari universitas dan institusi ternama. Misalnya, program *Google Data Analytics Certificate* di

Coursera sangat cocok bagi pemula, sementara MicroMasters *Data Science* dari MIT di edX menawarkan pembelajaran yang lebih mendalam. Di sisi lain, kursus Udemy sering kali lebih terjangkau dan langsung aplikatif.

Platform seperti Kaggle menyediakan lingkungan belajar berbasis proyek nyata dan kompetisi. Di sini, para pembelajar tidak hanya bisa mengakses dataset riil, tetapi juga belajar dari kode yang dibagikan oleh komunitas. Selain itu, komunitas profesional seperti *Towards Data Science* di Medium dan berbagai grup diskusi di LinkedIn, Reddit, maupun Discord, dapat menjadi sarana bertukar wawasan, mengembangkan jaringan, dan mengikuti tren terbaru di dunia data.

Penting juga untuk tidak sekadar menjadi konsumen informasi, tetapi mulai aktif berkontribusi. Menulis artikel, membagikan hasil analisis, atau membuat mini proyek akan memperkuat pemahaman dan membangun portofolio yang bermanfaat dalam pencarian kerja.

1.6.3 Tantangan dan Peluang di Masa Depan

Dunia *data analytics* bukan tanpa tantangan. Salah satu tantangan terbesar adalah isu privasi dan etika (Kaushik *et al.*, 2024). Ketika data semakin banyak dikumpulkan dari pengguna, muncul pertanyaan penting tentang siapa yang berhak mengakses, bagaimana data digunakan, dan bagaimana melindungi individu dari penyalahgunaan data. Di sinilah peran regulasi seperti GDPR (*General Data Protection Regulation*) di Eropa atau UU PDP (Pelindungan Data Pribadi) di Indonesia menjadi sangat penting sebagai

acuan. Selain itu, bias dalam algoritma juga menjadi tantangan tersendiri. Model prediktif yang dilatih dengan data historis bisa saja memperkuat ketimpangan sosial jika tidak dikontrol secara etis dan teknis. Namun, di balik tantangan ini terbuka peluang besar. Laporan *Future of Jobs Report 2025* dari *World Economic Forum* (WEF) menunjukkan bahwa pekerjaan di bidang data—termasuk *data analyst*, *data scientist*, dan *machine learning engineer*—merupakan pekerjaan dengan tingkat pertumbuhan tertinggi selama dekade ini (World Economic Forum, 2025). Dunia industri, pemerintahan, kesehatan, pendidikan, bahkan sektor pertanian dan energi kini sangat bergantung pada pengambilan keputusan berbasis data.

Mempelajari *data analytics* adalah investasi masa depan. Tidak hanya membuka peluang karier yang menjanjikan, tetapi juga memberikan keterampilan berpikir yang dibutuhkan dalam berbagai bidang kehidupan. Langkah awal memang menantang, tetapi dengan konsistensi belajar, rasa ingin tahu yang tinggi, dan keberanian mencoba, siapa pun dapat menjadi bagian dari revolusi data yang tengah berlangsung saat ini. Bab ini diakhiri dengan harapan bahwa pembaca tidak hanya memahami konsep dasar *data analytics*, tetapi juga merasa terdorong untuk mulai menjelajahnya lebih dalam—karena masa depan adalah milik mereka yang mampu membaca data dan bertindak atas dasar informasi yang tepat.

DAFTAR PUSTAKA

- Alkan, N., Menguc, K. and Kabak, Ö. (2022) 'Prescriptive Analytics: Optimization and Modeling', in *Springer Series in Advanced Manufacturing*, pp. 239–264. Available at: https://doi.org/10.1007/978-3-030-93823-9_9.
- Bhatlawande, S., Rajandekar, R. and Shilaskar, S. (2024) 'Implementing Middleware Architecture for Automated Data Pipeline over Cloud Technologies', in *Proceedings - 2024 13th IEEE International Conference on Communication Systems and Network Technologies, CSNT 2024*, pp. 506–513. Available at: <https://doi.org/10.1109/CSNT60213.2024.10546160>.
- Budiasto, J., Tallulembang, T.M., et al. (2025) *Machine Learning untuk Pemula: Konsep dan Implementasi*. Edited by T.Y. Zalni. Jakarta Utara: Penerbit Buku Indonesia.
- Budiasto, J., Ismanto, H., et al. (2025) *Python untuk Data Science: Panduan Praktis Menguasai Analisis Data*. Edited by W. Yuliani. Padang: Literasi Langsung Terbit.
- Chen, L., Mislove, A. and Wilson, C. (2015) 'Peeking Beneath the Hood of Uber', in *Proceedings of the 2015 Internet Measurement Conference*. New York, NY, USA: Association for Computing Machinery (IMC '15), pp. 495–508. Available at: <https://doi.org/10.1145/2815675.2815681>.
- Choudha, S. et al. (2024) 'Weather Prediction from Historical Data using Big Data Technologies', in *Proceedings of 5th International Conference on IoT Based Control Networks*

and Intelligent Systems, ICICNIS 2024, pp. 27–31. Available at: <https://doi.org/10.1109/ICICNIS64247.2024.10823212>.

Davenport, T.H. and Harris, J.G. (2007) *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press. Available at: <https://books.google.co.id/books?id=7NJDLKltepsC>.

Dhaouadi, J. *et al.* (2019) 'On the Use of Hyperparameter Optimization in Big Data Processing Pipelines: A Case Study', in *Proceedings - 2019 Innovations in Intelligent Systems and Applications Conference, ASYU 2019*. Available at: <https://doi.org/10.1109/ASYU48272.2019.8946352>.

Fu, Y. and Soman, C. (2021) 'Real-time Data Infrastructure at Uber', in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 2503–2516. Available at: <https://doi.org/10.1145/3448016.3457552>.

Geetha, V. and Sujatha, N. (2024) 'An Overview of Descriptive Analytics and Data Visualization', in *Proceedings of the 5th International Conference on Smart Electronics and Communication, ICOSEC 2024*, pp. 1158–1163. Available at: <https://doi.org/10.1109/ICOSEC61587.2024.10722273>.

Gomez-Uribe, C.A. and Hunt, N. (2016) 'The Netflix Recommender System: Algorithms, Business Value, and Innovation', *ACM Trans. Manage. Inf. Syst.*, 6(4). Available at: <https://doi.org/10.1145/2843948>.

- Gupta, A., Mehrotra, A. and Khan, P.M. (2015) 'Challenges of cloud computing & big data analytics', in *2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015*, pp. 1112–1115. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84960863160&partnerID=40&md5=d1077a9f11c1904230e4af4e96b7059b>.
- Igulu, K.T., Osuigbo, E. and Singh, T.P. (2023) 'Data Analytics in Business Intelligence', in *AI-Based Data Analytics: Applications for Business Management*, pp. 107–124. Available at: <https://doi.org/10.1201/9781032614083-7>.
- Joshi, A., Spilbergs, A. and Mikelsone, E. (2025) 'The Role of Digitalization, Big Data Analytics, and Artificial Intelligence in Transforming', *Journal of Neonatal Surgery*, 14(1S), pp. 1127–1135. Available at: <https://doi.org/10.52783/jns.v14.1686>.
- Jurnaidi, A. *et al.* (2025) *DATABASE*. Padang: Get Press Indonesia.
- Kaushik, P. *et al.* (2024) 'Ethical Considerations in Data Mining and Analytics', in *Proceedings - IEEE 2024 1st International Conference on Advances in Computing, Communication and Networking, ICAC2N 2024*, pp. 1516–1521. Available at: <https://doi.org/10.1109/ICAC2N63387.2024.10895012>.
- Kim, R.Y. (2023) 'Data-Driven User Experience Design', *Interactions*, 30(4), pp. 56–58. Available at: <https://doi.org/10.1145/3603717>.

- Krishnamoorthy, V. *et al.* (2024) 'Harnessing smart digital business models through advanced business analytics: Driving innovation and efficiency in the digital age', in *Cutting-Edge Technologies for Business Sectors*, pp. 57–85. Available at: <https://doi.org/10.4018/979-8-3693-9586-8.ch003>.
- Kropf, S.L. (2024) 'Big Data Analytics in Management Reporting: A Systematic Literature Review', in *Springer Proceedings in Business and Economics*, pp. 537–558. Available at: https://doi.org/10.1007/978-3-031-49105-4_32.
- Kumar, V. and L., M. (2018) 'Predictive Analytics: A Review of Trends and Techniques', *International Journal of Computer Applications*, p. Available at: <https://doi.org/10.5120/IJCA2018917434>.
- Kusumaningrum, D., Kurniati, N. and Santosa, B. (2021) 'Machine learning for predictive maintenance', in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 2348–2356. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85121142991&partnerID=40&md5=b69639b194e5df4fb12eb18de1c37571>.
- Lepenioti, K. *et al.* (2020) 'Prescriptive analytics: Literature review and research challenges', *Int. J. Inf. Manag.*, 50, pp. 57–70. Available at: <https://doi.org/10.1016/J.IJINFOMGT.2019.04.003>.
- Levendis, J. and Indika, N. (2025) 'The business analytics skills employers are looking for', *Decision Sciences*

- Journal of Innovative Education*, 23(2). Available at: <https://doi.org/10.1111/dsji.70004>.
- Mackenzie, I., Meyer, C. and Noble, S. (2013) *How retailers can keep up with consumers*. Available at: <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers> (Accessed: 20 April 2025).
- Marr, B. (2016) *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*. Wiley. Available at: <https://books.google.co.id/books?id=wAOmCgAAQBAJ>.
- Mayer-Schönberger, V. and Cukier, K. (2013) *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Pham, V. and Wan, Y. (2023) 'The Limit of Netflix's Digital Strategy on the Next Hit', in *29th Annual Americas Conference on Information Systems, AMCIS 2023*. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85192895015&partnerID=40&md5=ecf90df26b7c835dede1db2669329c90>.
- Provost, F. and Fawcett, T. (2013) *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O'Reilly Media. Available at: <https://books.google.co.id/books?id=4ZctAAAAQBAJ>.
- Saleh, H.H. *et al.* (2024) 'Enhancing Business Operations Efficiency thorough Predictive Analytics', *Journal of Ecohumanism*, p. Available at: <https://doi.org/10.62754/joe.v3i5.3932>.

- Sharma, S. *et al.* (2023) 'Hybrid Product Recommendation System using Popularity Based and Content-Based Filtering', in *2023 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2023*. Available at: <https://doi.org/10.1109/ICDSAAI59313.2023.10452564>.
- Siegel, E. (2013) *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, Or Die*. Wiley. Available at: <https://books.google.co.id/books?id=dGgs3JKUFQIC>.
- Singh, G., Singh, J. and Prabha, C. (2022) 'Data Visualization and its Key Fundamentals: A Comprehensive Survey', in *7th International Conference on Communication and Electronics Systems, ICCES 2022 - Proceedings*, pp. 1710–1714. Available at: <https://doi.org/10.1109/ICCES54183.2022.9835803>.
- Wang, H. *et al.* (2020) 'Prediction of breast cancer distant recurrence using natural language processing and knowledge-guided convolutional neural network', *Artificial Intelligence in Medicine*, 110. Available at: <https://doi.org/10.1016/j.artmed.2020.101977>.
- World Economic Forum (2025) *Future of Jobs Report 2025*. World Economic Forum. Available at: https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf (Accessed: 21 April 2025).
- Wu, Q. (2022) 'Real-time Predictive Analysis of Loan Risk with Intelligent Monitoring and Machine Learning Technique', in *2022 IEEE 4th International Conference on Power, Intelligent Computing and Systems, ICPICS 2022*, pp. 852–856. Available at: <https://doi.org/10.1109/ICPICS55264.2022.9873618>.

Zhang, J., Le, T. and Chen, J. (2023) 'Investigation of essential skills for data analysts: An analysis based on linkedin', *Journal of Global Information Management*, 31(1). Available at: <https://doi.org/10.4018/JGIM.326548>.

BAB 2

SUMBER DAN PENGOLAHAN DATA

Oleh Dodi

2.1 Pengantar Data

Data merupakan kumpulan objek yang bersifat fakta yang perlu dianalisis untuk mendapatkan sebuah kesimpulan. Data memiliki peran penting dalam berbagai bidang. Data dapat diklasifikasikan berdasarkan beberapa kategori utama:

a. Berdasarkan Sifatnya

Berdasarkan sifatnya, data dibagi menjadi 2, yaitu data kualitatif dan data kuantitatif. Data kualitatif adalah data yang berbentuk kategori atau deskriptif dan bukan berupa angka. Contoh: status pernikahan, tingkat kepuasan masyarakat. Sedangkan data kuantitatif adalah data yang berbentuk angka serta bisa diukur. Misalkan: jumlah tumbuhan, jumlah orang, nilai mata kuliah.

b. Berdasarkan Skala Pengukuran

Berdasarkan skala pengukuran, data dibagi menjadi data nominal, ordinal, interval dan rasio. Data nominal adalah data yang tidak memiliki urutan tertentu. Sedangkan data ordinal, adalah data yang

memiliki urutan atau tingkatan dengan ukuran yang sudah jelas.

Ciri-ciri data nominal adalah: berupa kategori atau label, tidak memiliki urutan atau tingkatan, tidak bisa dibandingkan secara numerik, serta bisa digunakan dalam operasi pengelompokan (*grouping*), tetapi tidak untuk operasi matematika. Contohnya adalah jenis kelamin laki-laki dan perempuan; warna favorit adalah kuning, ungu, putih; provinsi tempat tinggal: Kalimantan Barat, Jakarta, Sumatera Barat; jenis pekerjaan Nelayan, Guru, Mekanik; nomor punggung pemain sepak bola, yang merupakan angka namun hanya sebagai identifikasi.

Ciri-ciri data ordinal adalah memiliki kategori atau label; Memiliki urutan atau tingkatan peringkat lebih tinggi berarti lebih baik, namun tanpa jarak yang jelas. Contoh Data Ordinal adalah: peringkat akademik yang terdiri atas rendah, sedang, dan tinggi; tingkat kepuasan konsumen, sangat puas, puas, tidak puas; peringkat lomba, Juara 1, Juara 2, Juara 3; tingkat pendidikan, SD, SMP, SMA, S1, S2, S3.

Data interval adalah data numerik dengan jarak yang sama antar nilai dan tidak memiliki nol mutlak. Sedangkan data rasio adalah data numerik dengan jarak yang sama dan memiliki nol mutlak.

Ciri-Ciri Data Interval:

- Urutan data yang jelas, data dapat disusun dari yang paling kecil hingga ke data yang nilainya paling besar.
- jarak antar data adalah nilainya sama, jarak antara 10 dan 20 sama dengan jarak antara 30 dan 40.
- Tidak memiliki nol absolut (nol bukan berarti tidak ada), nol dipandang sebagai titik referensi, bukan sebagai ketiadaan nilai.
- Data dapat digunakan untuk operasi matematika tertentu.

Contoh Data Interval:

- Suhu dalam Celsius atau Fahrenheit: 0°C tidak berarti "tidak ada suhu" (berbeda dengan nol dalam skala Kelvin); Perbedaan antara 20°C dan 30°C sama dengan perbedaan antara 30°C dan 40°C .
- Skor IQ: Skor IQ 0 tidak berarti seseorang tidak memiliki kecerdasan; Jarak antara 100 dan 110 sama dengan jarak antara 120 dan 130.
- Tahun dalam kalender: tahun 0 Masehi hanyalah titik referensi; tidak berarti tidak ada waktu sebelum itu.

Ciri-Ciri Data Rasio:

- Memiliki urutan yang jelas, data dapat disusun dari data dengan nilai terkecil ke data dengan nilai terbesar.
- Jarak antar nilai sama, perbedaan antara 10 dan 20 sama dengan perbedaan antara 30 dan 40.
- Memiliki nol mutlak, nilai nol menunjukkan ketiadaan nilai, bukan hanya titik referensi.
- Dapat digunakan dalam semua operasi matematika.

Contoh Data Rasio:

- Berat badan, 0 kg berarti tidak ada berat sama sekali; 60 kg adalah tiga kali lebih berat dari 20 kg.
- Pendapatan: jika seseorang memiliki Rp 0, itu berarti tidak memiliki uang sama sekali.
- Panjang atau tinggi, 0 cm berarti tidak ada Panjang; 300 cm adalah tiga kali lebih panjang dari 100 cm.
- Jumlah siswa dalam kelas: tidak bisa ada -5 siswa; nol adalah batas absolut.
- Waktu: 0 detik berarti tidak ada waktu yang berlalu.

2.2 Sumber Data

2.1.1 Data Primer dan Data Sekunder

Berdasarkan sumbernya, data dibagi menjadi data primer dan data sekunder. Data Primer merupakan data yang dikumpulkan secara langsung oleh peneliti melalui survei, wawancara, atau eksperimen. Sedangkan data sekunder merupakan data yang diperoleh dari sumber lainnya seperti data-data yang sudah disediakan oleh lembaga terpercaya.

Data primer merupakan data yang dikumpulkan langsung oleh peneliti dari sumber asli. Tentu data dari sumber asli lebih akurat dan spesifik serta relevan dan valid. Ciri-ciri data primer adalah: dikumpulkan langsung dari sumber pertama, lebih akurat karena dikumpulkan dengan tujuan spesifik, membutuhkan waktu, biaya, dan tenaga lebih besar dibandingkan data sekunder, dan pengumpulan datanya dapat menggunakan berbagai metode pengumpulan data seperti survei, wawancara, observasi, dan eksperimen. Contoh: wawancara dengan Rektor atau Direktur tentang strategi peningkatan mutu perguruan tinggi.

Data sekunder merupakan data yang sudah dikumpulkan dan dipublikasi oleh pihak lain, sebelum digunakan kembali oleh pihak lain. Pihak pengguna data sekunder tidak harus mengumpulkan data dari awal. Data sekunder dapat diperoleh dari internal dan eksternal. Sumber internal dapat berasal dari dalam organisasi atau institusi tertentu, seperti laporan keuangan, data pelanggan, arsip data dari perusahaan. Sumber Eksternal dapat diperoleh dari lembaga-lembaga tertentu, seperti laporan pemerintah, data dari

badan statistik, jurnal ilmiah, hasil survei sebelumnya, media massa, serta database dari lembaga penelitian atau akademik. Keunggulan data sekunder adalah efisien terhadap penggunaan waktu dan biaya, tersedia dalam bentuk digital maupun cetak, cakupan data yang luas secara historis dan wilayah.

2.1.2 Validitas Data

Validitas data adalah sejauh mana data yang dikumpulkan benar-benar mengukur apa yang seharusnya diukur dan dapat dipercaya untuk digunakan dalam analisis. Kevalidan suatu data menentukan keakuratan hasil analisis dan relevan dengan tujuan penelitian. Sedangkan data yang tidak valid menyebabkan kesalahan dalam penarikan suatu kesimpulan dan berakibat biasanya hasil penelitian. Oleh karena itu, setiap instrumen pengumpulan data, seperti kuesioner, wawancara, atau eksperimen, harus diuji validitasnya sebelum digunakan.

Berdasarkan cara pengukurannya validitas data dapat dibedakan menjadi beberapa jenis, yaitu: validitas konstruk, validitas isi, validitas konkuren, dan validitas eksternal. Validitas konstruk adalah validitas yang digunakan untuk menilai apakah instrumen penelitian benar-benar mengukur variabel akan diteliti. Contoh: Jika kuesioner digunakan untuk mengukur kepuasan konsumen, pertanyaannya harus benar-benar mencerminkan faktor-faktor kepuasan, bukan hal lain seperti aktifitas atau pendapatan konsumen. Validitas isi merupakan validitas yang mencakup untuk mengukur sejauh mana instrumen mencakup seluruh aspek dari konsep yang diteliti. Contoh: tes mata kuliah statistik

yang valid harus mencakup semua topik penting dalam kurikulum, bukan hanya sebagian kecil materi saja.

Validitas prediktif, digunakan untuk mengukur kemampuan suatu data dalam memprediksi hasil di masa depan. Contoh: Skor ujian masuk perguruan tinggi seharusnya bisa memprediksi seberapa baik mahasiswa akan berprestasi selama kuliah. Validitas konkuren digunakan untuk membandingkan hasil suatu alat ukur dengan alat ukur lain yang sudah terbukti valid pada waktu bersamaan. Contoh: jika dua tes kecerdasan dilakukan secara bersamaan dan memberikan hasil yang mirip, maka keduanya memiliki validitas konkuren. Validitas eksternal digunakan untuk mengukur apakah hasil penelitian dapat digeneralisasikan ke populasi yang lebih luas. Misalkan: jika hasil survei konsumen dari 10 responden dapat mewakili seluruh konsumen suatu perusahaan, maka validitas eksternalnya tinggi. Oleh karena itu, validitas data merupakan hal yang sangat penting karena: menghindari kesalahan analisis data, meningkatkan kepercayaan terhadap penelitian yang dilakukan, dan mendukung pengambilan keputusan yang lebih baik.

2.1.3 Reliabilitas Data

Reliabilitas data adalah tingkat konsistensi terhadap suatu instrumen pengukuran yang akan digunakan dalam penelitian untuk menghasilkan data yang sama apabila digunakan berulang kali pada kondisi yang sama. Apabila instrumen yang digunakan menghasilkan nilai yang konsisten maka instrumen tersebut dianggap reliabel.

Reliabilitas dapat dibedakan menjadi beberapa jenis, antara lain: reliabilitas uji-ulang, reliabilitas konsistensi internal, reliabilitas antar-penilai, dan reliabilitas bentuk sejajar. Reliabilitas uji-ulang digunakan untuk mengukur sejauh mana hasil yang diperoleh tetap konsisten ketika tes atau instrumen yang sama digunakan kembali pada waktu yang sama. Contoh: Jika seseorang mengisi kuesioner kepuasan responden hari ini dan mengisi ulang bulan depan dengan hasil yang mirip, maka reliabilitasnya tinggi.

Reliabilitas konsistensi internal digunakan untuk mengukur sejauh mana butir pertanyaan pada sebuah instrumen saling berkorelasi atau konsisten dalam mengukur objek yang sama. Contoh: Jika dalam survei kepuasan konsumen, ada 10 pertanyaan yang mengukur kepuasan, maka setiap pertanyaan harus memberikan hasil yang sejalan. Reliabilitas antar-penilai digunakan untuk mengukur tingkat kesepakatan antara dua atau lebih penilai dalam memberikan skor atau menilai suatu hal. Contoh: jika dua dosen menilai esai mahasiswa dengan hasil yang hampir sama, berarti tingkat reliabilitas antar-penilai tinggi. Reliabilitas bentuk sejajar digunakan untuk mengukur sejauh mana dua versi instrumen yang berbeda tetapi setara memberikan hasil yang sama. Contoh: Jika ada dua versi tes kemampuan akademik yang berbeda tetapi menghasilkan skor yang mirip, berarti instrumen tersebut reliabel.

Mengukur suatu reliabilitas data sangat penting karena: menjamin keakuratan penelitian hingga menghasilkan analisis yang lebih dapat dipercaya, mencegah kesalahan interpretasi karena hasil yang tidak konsisten dapat menyebabkan kesimpulan yang salah,

serta memastikan replikasi penelitian, karena jika penelitian diulang dengan metode yang sama, hasilnya seharusnya tetap konsisten.

2.1.4 Etika dalam Pengumpulan Data

Etika dalam pengumpulan data merupakan hal yang sangat penting untuk diperhatikan. Etika memastikan bahwa proses penelitian dilakukan secara adil, transparan, dan menghormati hak-hak orang lain. Etika yang diperhatikan dalam pengumpulan data diantaranya adalah persetujuan responden. Responden harus diberi informasi yang jelas tentang tujuan penelitian, bagaimana data mereka akan digunakan, dan responden diberikan keleluasaan untuk menolak dan menghentikan partisipasinya tanpa ada konsekuensi. Responden tidak dipaksa untuk memberikan data.

Etika yang perlu diperhatikan selanjutnya adalah kerahasiaan dan privasi data responden. Data pribadi responden harus dijaga kerahasiaannya. Data responden hanya dapat digunakan untuk tujuan penelitian yang telah direncanakan sejak awal. Hal ini juga bertujuan untuk meminimalisir resiko bagi responden secara fisik, emosional, atau psikologis. Jika ada risiko yang harus dihadapi oleh responden, harus diberi tahu sebelumnya.

Dalam pengumpulan data, peneliti hendaknya memilih responden secara adil, tanpa diskriminasi atau bias, dan tidak boleh ada paksaan dalam berpartisipasi. Data yang dikumpulkan hanya boleh digunakan sesuai dengan tujuan yang telah disetujui, dan tidak boleh disalahgunakan untuk kepentingan lain yang dapat merugikan responden. Oleh karena itu, peneliti harus mengikuti regulasi dan pedoman etika penelitian yang

berlaku di wilayah atau institusi tempat penelitian dilakukan.

2.3 Pengolahan Data

a. Pembersihan Data

Pembersihan data adalah proses mencari dan memperbaiki kesalahan pada sebuah data sebelum digunakan untuk dianalisis. Tujuannya adalah memastikan bahwa data yang digunakan berkualitas tinggi, dan akurat. Dalam statistik, data yang dikumpulkan sering kali mengandung kesalahan seperti duplikasi atau terdapat data yang tidak valid. Jika dibiarkan, kesalahan ini dapat menyebabkan proses analisis data yang tidak bagus.

Misalkan, pada nilai ujian mata kuliah statistik, nilai seorang mahasiswa tercatat tiga kali dalam penyimpanan. Oleh karena itu, diperlukan adanya pembersihan data terhadap data-data yang sama. Kasus lain yang sering muncul adalah adanya data yang hilang dikarenakan datanya tidak ditemukan atau tidak diberikan oleh responden. Oleh karena itu, peneliti harus mengganti data yang hilang dengan metode tertentu agar data tersebut valid.

b. Transformasi dan Normalisasi Data

Transformasi data adalah proses mengubah data agar lebih sesuai untuk analisis. Tujuannya adalah mengatasi skala yang tidak seragam dalam dataset, mengurangi dampak pencilan (*outlier*), membantu memenuhi asumsi statistik tertentu (misalnya, normalitas dalam regresi), serta untuk

meningkatkan interpretabilitas data. Sedangkan normalisasi data adalah proses mengubah nilai-nilai pada suatu data agar berada dalam skala yang sama tanpa mengubah pola distribusinya. Normalisasi sering digunakan untuk memastikan bahwa setiap fitur memiliki pengaruh yang seimbang dalam analisis. Tujuan dilakukannya normalisasi data adalah memastikan bahwa data memiliki skala yang seragam, menghindari dominasi fitur dengan nilai besar terhadap fitur dengan nilai kecil, serta memudahkan algoritma statistik untuk bekerja lebih optimal.

Terdapat berbagai jenis transformasi data, yaitu:

1. Transformasi Logaritma

Transformasi logaritma menggunakan nilai logaritma untuk mengurangi skala data yang sangat bervariasi dan membuat distribusinya lebih normal. Biasanya digunakan ketika ada perbedaan yang cukup tinggi antara nilai minimum dan maksimum. Misalkan, pada data pendapatan responden berkisar antara dari Rp 1.00.000 hingga Rp1.000.000.000, kita bisa menerapkan transformasi logaritma agar distribusinya lebih merata. Persamaan yang digunakan adalah:

$$X' = \log X$$

X' merupakan nilai pada data yang baru.

2. Transformasi Akar Kuadrat

Transformasi akar kuadrat berguna untuk mengurangi variasi data dengan tetap mempertahankan hubungan proporsional antar data. Transformasi ini dapat diterapkan

pada data yang tidak memiliki angka nol atau negatif. Misalkan, apabila data tentang jumlah konsumen dalam sehari memiliki distribusi yang terlalu miring, maka dapat menggunakan transformasi akar kuadrat untuk menyeimbangkannya. Persamaan yang digunakan adalah:

$$X' = \sqrt{X}$$

X' merupakan nilai pada data yang baru.

Terdapat berbagai cara untuk normalisasi data, diantaranya adalah:

1. *Min-Max Scaling*

Cara *Min-Max Scaling* mengubah nilai data ke dalam rentang tertentu, biasanya antara 0 dan 1. Rumus yang digunakan adalah:

$$X' = \frac{X - X_{min}}{X_{maks} - X_{min}}$$

X' merupakan nilai pada data yang baru.

Misalkan, pada sebuah kasus, data skor ujian antara 50 hingga 100. Data setelah dinormalisasi, berada dalam rentang 0 hingga 1. Data sebelum normalisasi: 50, 60, 80, 100. Data setelah normalisasi: 0, 0.2, 0.6, 1

2. *Z-Score Normalization*

Cara ini *Z-Score Normalization* yaitu mengubah data sehingga memiliki rata-rata 0 dan standar deviasi 1. Rumus yang digunakan adalah:

$$X' = \frac{X - \mu}{\sigma}$$

X' merupakan nilai pada data yang baru.

μ adalah rata-rata data

σ adalah standar deviasi data

Misalkan, jika tinggi badan mahasiswa memiliki rata-rata 165 cm dan standar deviasi 10 cm, maka seorang siswa dengan tinggi 175 cm akan memiliki nilai standar 1.

3. *Decimal Scaling*

Cara *Decimal Scaling* yaitu mengubah data dengan membagi setiap nilai dengan kelipatan 10 yang sesuai dengan jumlah digit terbesar. Rumusnya adalah:

$$X' = \frac{X}{10^j}$$

X' merupakan nilai pada data yang baru.

j adalah jumlah digit dari angka terbesar pada data

Misalkan terdapat data tertinggi adalah 885, maka $j=3$, dan semua nilai dibagi dengan 1000.

c. Pengkodean dan Kategorisasi Data

Pengkodean data adalah proses mengubah data yang bukan angka menjadi format angka agar bisa digunakan dalam analisis statistik. Sedangkan kategorisasi data adalah proses pengelompokkan data berdasarkan karakteristik tertentu agar lebih mudah dianalisis.

Pengkodean dapat dilakukan dengan berbagai cara, yaitu *label encoding*, *one-hot encoding*, *binary encoding*, dan *frequency encoding*. Penjelasan sebagai berikut:

1. *Label Encoding*

Label Encoding dengan cara mengubah kategori menjadi angka berdasarkan urutan

tertentu. Misalkan data pada jenjang pendidikan TK = 0, SD = 1, SMP = 2, SMA = 3.

2. *One-Hot Encoding*

One-Hot Encoding mengubah kategori menjadi variabel biner yaitu 0 dan 1. Misalkan pada data warna:

Kuning adalah [1, 0, 0]

Hijau adalah [0, 1, 0]

Putih adalah [0, 0, 1]

3. *Binary Encoding*

Binary Encoding mengubah kategori menjadi representasi biner. Misalkan kita memiliki 4 kategori: A adalah 00, B adalah 01, C adalah 11, dan D adalah 10.

4. *Frequency Encoding*

Frequency Encoding mengubah kategori menjadi angka berdasarkan frekuensi kemunculan.

Kategorisasi dapat dilakukan dengan berbagai cara, yaitu berdasarkan rentang nilai, berdasarkan aturan atau domain keilmuan, dan berdasarkan *clustering*. Penjelasannya sebagai berikut:

1. Kategorisasi Berdasarkan Rentang Nilai

Kategorisasi ini mengelompokkan data numerik ke dalam beberapa kategori. Misalkan pada rentang 0-12 tahun dikelompokkan ke dalam usia anak-anak, rentang 13-18 tahun berada pada usia remaja, rentang pada usia 19-60 tahun berada pada usia dewasa, dan usia diatas 60 tahun dikeompokkan ke dalam usisa lansia.

2. Kategorisasi Berdasarkan Aturan atau Domain Keilmuan

Kategorisasi ini mengelompokkan data berdasarkan pengetahuan atau standar yang ada. Misalkan penilaian pada sistem pendidikan, nilai 0 – 50 berada pada kategori Kurang, 51 – 70 dianggap Cukup, 71 – 85 dianggap Baik, dan nilai 86 – 100 dianggap Sangat Baik.

3. Kategorisasi Berdasarkan *Clustering*

Kategorisasi ini menggunakan teknik statistik seperti K-Means Clustering untuk mengelompokkan data ke dalam kategori berdasarkan pola dalam dataset. Misalkan: analisis konsumen, dapat dikelompokkan konsumen berdasarkan waktu belanja.

d. Pengolahan Data dengan Perangkat Lunak

Perangkat lunak merupakan media yang dapat digunakan untuk membantu pengolahan data. Perangkat lunak membantu dalam berbagai tugas, seperti pembersihan data, transformasi, visualisasi, dan pembuatan laporan. Banyak jenis perangkat lunak yang dapat digunakan untuk pengolahan data, misalkan perangkat lunak spreadsheet, perangkat lunak statistik, perangkat lunak pemrograman.

Perangkat lunak Spreadsheet digunakan untuk mengolah data dalam bentuk tabel dan melakukan analisis dasar. Misalkan microsoft excel, dapat digunakan untuk perhitungan, pengolahan data, dan visualisasi. Google Sheets dapat digunakan sebagai alternatif berbasis cloud dengan fitur kolaborasi. Perangkat lunak statistik dapat

digunakan untuk analisis statistik lanjutan, seperti regresi, uji hipotesis, dan pemodelan data. Misalkan, aplikasi SPSS dapat digunakan dalam penelitian akademik dan sosial untuk analisis statistik kompleks. Perangkat lunak pemrograman untuk pengolahan data dapat digunakan untuk manipulasi data yang lebih fleksibel, terutama dalam analisis data besar (*big data*). Misalkan *python* (dengan pustaka seperti Pandas, NumPy, dan Matplotlib) dapat digunakan dalam statistik dan analisis data ilmiah.

Selain itu, terdapat juga perangkat lunak yang digunakan untuk visualisasi data. Perangkat lunak tersebut digunakan untuk membuat grafik dan dashboard interaktif dari data. Misalkan tableau, merupakan perangkat lunak visualisasi data yang kuat dan interaktif. Google Data Studio sebagai perangkat lunak alternatif berbasis cloud untuk membuat laporan interaktif.

DAFTAR PUSTAKA

- Arifin dan Aunillah. (2021). *Buku Ajar: Statistik Pendidikan*. Sidoarjo: Umsida Press.
- Arikunto, S. (2018). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Cooper, D.R., & Schindler, P.S. (2014). *Business Research Methods (12th ed.)*. McGraw-Hill Education.
- Creswell, J.W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (4th ed.)*. SAGE Publications.
- Nursalam. (2015). *Statistika Pendidikan*. Makassar: Alauddin University Press.
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Wahjusaputri, Sinta dan Purwanto, Anim. (2022). *Statistika Pendidikan, Teori dan Aplikasi*. Yogyakarta: Percetakan Bintang.
- Wahyuni, Molli. (2020). *Statistik Deskriptif untuk Penelitian Olah Data Manual dan SPSS Versi 25*. Yogyakarta: Bintang Pustaka Madani.

BAB 3

TEKNIK ANALISIS DATA

Oleh Ketrin R. Manullang

3.1 Pendahuluan

Dalam sebuah penelitian, pengumpulan semua data pada penelitian bertujuan untuk menarik sebuah kesimpulan dalam penelitian tersebut, oleh karena itu pasti mengolah dan menganalisis data perlu dilakukan. Proses pengolahan dan analisis data ini perlu dilakukan karena data mentah yang diperoleh dalam pengumpulan data penelitian tidak bisa secara langsung dan eksplisit dapat ditarik kesimpulannya. Diperlukan sebuah sistem tata cara dalam mengkonstruksi teori, membangun hipotesis, dan menguji hipotesis sebuah penelitian. Dan sistem tersebut adalah teknik analisis data.

Teknik analisis data yang diimplementasikan dalam sebuah penelitian harus ditentukan dan dijelaskan secara tepat. Teknik analisis yang tepat maka peneliti dapat memahami dan memaknai makna data secara mendalam. Selain itu peneliti juga dapat memahami hubungan dan pola yang dimiliki data dan antar data. Hal ini perlu dilakukan agar teknik analisis yang diaplikasikan dapat membuat generalisasi dan kesimpulan sebuah penelitian.

3.2 Definisi Teknik Analisis Data

Kegiatan yang dilaksanakan bertujuan menetapkan makna yang sebenarnya berdasarkan data yang telah dikelola dalam bentuk pola yang logis disebut analisis data (Fadjarajani, dkk, 2020:11). Analisis data diterapkan untuk menjawab masalah penelitian (Mukhid, 2021:17). Analisis data meliputi pengurutan dan penyederhanaan data (Mukhid, 2021:181). Penyederhanaan data ini dilakukan agar data-data dalam penelitian dapat dikaitkan langsung dengan masalah penelitian atau kajian penelitian yang sedang dikaji.

Analysis berasal dari bahasa Yunani, yaitu “ana” yang bermakna atas dan “lysis” yang bermakna memecahkan (Kurniasih, dkk, 2021:5). Analisis berarti pemisahan atau pemeriksaan yang teliti. Proses analisis dilakukan terhadap data-data penelitian yang sudah dikumpulkan. Analisis berguna untuk memahami dan memaknai data hasil penelitian. Analisa data dikerjakan untuk memvalidasi atau membuktikan hipotesa (Amelia, 2023:155).

Analisis data berarti jalan mengukur rangkaian data, menyusunnya ke dalam suatu model, kelompok dan penjabaran dasar (Nurdewi, 2022). Teknik analisis data dapat diartikan sebagai suatu prosedur mengorganisasikan susunan data, mengaturnya ke dalam suatu pola, kelompok dan unit dasar (Rahmadi, 2011:92). Pada tahap analisis ini dapat mengaplikasikan teknik statistik yang sesuai dengan tujuan penelitian (Widodo, dkk, 2023:41). Statistik digunakan untuk menyederhanakan data penelitian yang jumlahnya banyak sehingga mudah untuk dipahami informasinya.

Analisis data di penelitian kuantitatif menggunakan metode yang bermacam-macam dengan mengaplikasikan statistik. Sedangkan dalam penelitian kualitatif, analisis data menggunakan metode dengan pengkodean (*coding*) dan pengkategorian data dari kajian yang diteliti dalam sebuah penelitian. Jadi sangat jelas bahwa jenis penelitian akan menentukan teknik yang digunakan dalam melakukan analisis data.

3.3 Langkah-langkah Teknik Analisis Data

Tahapan dalam melakukan analisis data dimulai dengan melaksanakan analisa deskriptif dan kemudian dilanjutkan dengan analisa lanjut (statistik inferensial). Uraian tentang langkah-langkah analisis dapat diuraikan sebagai berikut.

1. Analisa Deskriptif

Menurut Sugiyono (dalam Abdullah, dkk, 2022:88) teknik analisis statistik deskriptif merupakan metode analisis data dengan memaparkan data, tanpa menarik kesimpulan yang diterapkan untuk populasi (generalisasi). Deskriptif berarti menjelaskan sedetail mungkin tentang realitas yang dikaji (Haryoko, Bahartiar, dan Arwadi, 2020:16). Analisa deskriptif bertujuan untuk memberikan gambaran umum data penelitian.

Metode ini biasanya hanya bermaksud untuk mendeskripsikan data penelitian, tapi tidak bertujuan untuk membuat kesimpulan data penelitian (Abdullah, dkk, 2022:90). Analisa deskriptif menganalisis data secara sistematis.

2. Analisa Lanjut (Statistik inferensial)

Metode statistik inferensial ini membahas metode menarik kesimpulan mengenai karakteristik populasi dengan memakai data sampel yang terbatas (Ibrahim, dkk, 2018:115). Statistik inferensial disebut juga statistik analitik karena analisis ini digunakan untuk membuat generalisasi hasil penelitian.

Untuk mengetahui antar kelompok secara deskriptif dalam teknik analisis statistik inferensial, diawali dengan menghitung nilai mean dan standar deviasi. Pengujian pada statistik inferensial yaitu Uji Normalitas, Uji Linieritas, dilanjutkan dengan Analisis Regresi Linier Sederhana, interpretasi hasil, penarikan kesimpulan dan kemudian pelaporan. Uji Normalitas termasuk dalam uji persyaratan analisis dalam statistik parametrik. Selain Uji Normalitas, juga terdapat beberapa uji lainnya sebagai uji persyaratan dalam statistik parametrik, yaitu Uji Homogenitas Variansi.

3.4 Jenis-Jenis Teknik Analisis Data

Teknik analisis data yang diterapkan dalam sebuah penelitian ditentukan oleh jenis penelitian yang diimplementasikan dalam sebuah penelitian. Jenis penelitian dikelompokkan menjadi 2 (dua), yaitu 1) penelitian kuantitatif, dan 2) penelitian kualitatif. Berdasarkan jenis penelitian ini, maka teknik analisis data dibagi menjadi teknik analisis data penelitian kuantitatif dan teknik analisis data penelitian kualitatif.

1. Teknik Analisis Data Penelitian Kuantitatif

Teknik analisis data kuantitatif seperti regresi, analisis varians, dan uji korelasi (Rachman, dkk, 2024:105). Teknik analisis data kuantitatif di penelitian kuantitatif menerapkan statistik (Abdullah, dkk, 2022:87). Uji statistik bisa seperti uji t, korelasi, analisa faktor, analisa varian dan covarian, dan regresi linear. Menurut Sahir (2021:37) teknik analisis data kuantitatif menerapkan 2 metode statistik. Kedua metode statistik tersebut adalah statistik deskriptif dan statistik inferensial.

a. Statistik Deskriptif

Statistik deskriptif biasanya sebatas pada pendeskripsian variabel penelitian satu per satu (Ibrahim, dkk, 2018:47). Statistik deskriptif digunakan untuk mendeskripsikan data sebagaimana adanya tanpa menghasilkan kesimpulan umum (generalisasi) (Mukhid, 2021:213). Analisa statistik deskriptif ini biasanya digunakan pada penelitian deskriptif. Istilah lain untuk metode statistik deskriptif yaitu statistika eksploratif, sedangkan untuk statistika inferensial yaitu statistik induktif atau konfirmasi (Haryono, 2012:91).

Penelitian deskriptif menggambarkan secara sistematis, tentang realitas dan karakteristik populasi (K, 2013:27). Penelitian deskriptif pada dasarnya menyelidiki variabel secara independen, seringkali hanya satu, tanpa dihubungkan dengan variabel lain (Rahmadi, 2011:27).

Analisis data statistik deskriptif membatasi generalisasinya pada kelompok individu tertentu yang diamati (Mukhid, 2021:181). Statistik

deskriptif membahas alat, teknik, dan prosedur yang bertujuan untuk menguraikan sekumpulan data dari hasil pengamatan (Panjaitan dan Ahmad, 2017:72).

b. Statistik Inferensial

Teknik analisis inferensial melihat keragaman hubungan antara variabel (Sahir, 2021:38). Statistika inferensial (induktif) bertujuan untuk menarik kesimpulan. Suatu dugaan hasil dari perhitungan statistika deskriptif diperlukan sebelum dilakukan penarikan kesimpulan (Purnomo, 2016:18). Hasil dari statistik inferensial akan digeneralisasikan untuk populasi, dan inilah merupakan tujuan penggunaan statistik inferensial.

Statistik inferensial dikelompokkan menjadi 2 (dua). Kedua jenis tersebut adalah 1) statistik parametrik, dan 2) statistik non parametrik. Menurut Purnomo (2016:16) metode statistik parametrik adalah dipakai untuk sebaran data normal, dan metode statistik non-parametrik tidak mensyaratkan sebaran data normal. Statistik parametrik bertujuan menguji hipotesis korelasional, kausatif, asosiatif, dan prediktif. Dan statistik non-parametrik bertujuan menguji hipotesis komparatif (Haryono, 2012:236).

Dalam statistik inferensial dilakukan upaya untuk menarik kesimpulan dan menarik keputusan berlandaskan analisis perhitungan statistik. Istilah lain untuk statistik inferensial ini disebut dengan istilah statistik deduktif. Hal ini karena dilakukan uji empiris teori dan pengukuran terhadap variabel, serta generalisasi hasil penelitian terhadap populasi.

2. Teknik Analisis Data Penelitian Kualitatif

Penjaringan data di penelitian kualitatif didapat melalui beragam metode pengumpulan data, dan dikerjakan secara berkali-kali hingga datanya jenuh (Sidiq dan Qoiri, 2019:38). Analisis data dilakukan secara induktif (Warsono, Astuti, dan Ardiyansah, 2022:11). Artinya kesimpulan diambil berdasar pada realitas yang ada di lapangan, tidak berdasar teori (Abdussamad, 2021:85).

Kegiatan analisis data dilakukan dari saat sebelum, memasuki, selama, dan setelah selesai di lapangan. Metode kualitatif berfokus pada interpretasi, pemahaman konteks, dan makna subjektif (Rachman, dkk, 2024:137). Analisis data dilakukan melalui prosedur pembentukan pola, tema, atau ciri umum (Haryoko, Bahartiar, dan Arwadi, 2020:23).

Analisis data dilakukan secara bersamaan atau simultan (Haryoko, Bahartiar, dan Arwadi, 2020:27). Analisis data menerapkan pendekatan teknik analisis kualitatif. Kegiatan analisis data ini bertujuan untuk memaparkan makna/arti yang dimiliki oleh data penelitian. Data-data ini biasanya bersifat dinamis dan prosesnya selalu berkembang dinamis.

Menurut Sugiyono (dalam Abdussamad, 2021:103) terdapat 3 (tiga) langkah dalam penelitian kualitatif, yaitu 1) deskripsi, 2) reduksi, dan 3) seleksi. Menurut Hardani, dkk (2020:21) metode penelitian kualitatif pada umumnya menerapkan metode yang berbentuk siklus.

Analisis data bersifat induktif, artinya analisis berdasarkan data yang didapat, kemudian dibangun menjadi hipotesis (Abdussamad, 2021:159). Menurut

Ilyas (2015:84), analisis data induktif adalah analisis data lapangan kedalam bagian-bagian dan dilanjutkan dengan pengelompokan.

3.4 Perbedaan Teknik Analisis Data Kuantitatif dan Teknik Analisis Data Kualitatif

Terdapat perbedaan dalam analisis data pada teknik analisis data kuantitatif dan teknik analisis data kualitatif. Perbedaan antara teknik analisis data kuantitatif dan teknik analisis data kualitatif adalah sebagai berikut.

1. Tujuan dilakukan analisis data dalam penelitian kualitatif adalah untuk membentuk hipotesis, sedangkan analisis data dalam penelitian kuantitatif dilakukan untuk menguji hipotesis (Abdussamad, 2021:81).
2. Analisis data dalam penelitian kualitatif dilakukan secara induktif sedangkan pada penelitian kuantitatif secara deduktif. Induktif yaitu pengambilan kesimpulan berdasar fakta-fakta di lapangan, tidak berdasar teori tertentu. Deduktif yaitu uji empiris terhadap teori yang ada dengan sarana statistik (Rukminingsih, Adnan, dan Latief, 2020:20).
3. Data yang dikumpulkan pada penelitian kuantitatif berupa angka, kesimpulan diambil berdasarkan hasil analisis statistik. Sedangkan data pada penelitian kualitatif dapat berupa verbal, dan analisis data tanpa menggunakan penghitungan teknik statistik.
4. Penelitian kualitatif disusun secara naratif yang kreatif dan mendalam, naturalistik yang otentik. Penelitian kuantitatif berfokus pada pengukuran dan analisis perhitungan variabel, bukan prosesnya, penyelidikan dinilai berada dalam kerangka netral yang tidak memihak (Hardani, dkk. 2020:254).

5. Hasil penelitian pada penelitian pada penelitian kualitatif tidak dapat digeneralisasik, sedangkan pada penelitian kualitatif hasil penelitian dapat dilakukan penarikan kesimpulan yang diaplikasikan untuk umum (generalisasi).

DAFTAR PUSTAKA

- Abdullah, Karimuddin, dkk. 2022. *Metodologi Penelitian Kuantitatif*. Kab. Pidie Provinsi Aceh: Yayasan Penerbit Muhammad Zaini.
- Abdussamad, Zuchri. 2021. *Metode Penelitian Kualitatif*. Makassar: CV. Syakir Media Press
- Amelia, Dahlia, dkk. 2023. *Metode Penelitian Kuantitatif*. Provinsi Aceh: Yayasan Penerbit Muhammad Zaini.
- Fadjarajani, Siti. dkk. 2020. *Metodologi Penelitian Pendekatan Multidispliner*. Kota Gorontalo: Ideas Publishing.
- Hardani, dkk. 2020. *Metode Penelitian Kualitatif & Kuantitatif*. Yogyakarta: Penerbit Pustaka Ilmu.
- Haryoko, Sappto, Bahartiar, dan Arwadi, Fajar. 2020. *Analisis Data Penelitian Kualitatif*. Makassar Badan Penerbit UNM.
- Haryono, Siswoyo. 2012. *Metodologi Penelitian Manajemen: Teori dan Aplikasi*. Jakarta: PT Intermedia Personalia Utama (IPU).
- Ibrahim, Andi. dkk. 2018. *Metodologi Penelitian*. Makassar: Gunadarma Ilmu.
- Ilyas, Muhammad. 2015. *Metodologi Penelitian Pendidikan Matematika*. Bandung: Pustaka Ramadhan.
- K, Abdullah. 2013. *Tahapan dan Langkah-langkah Penelitian*. Watampone: Luqman Al-Hakim Press
- Kurniasih, Dewi, dkk. 2021. *Teknik Analisa*. Bandung: CV Alfabeta.
- Mukhid, Abd. 2021. *Metodologi Penelitian Pendekatan Kuantitatif*. Surabaya: CV. Jakad Media Publishing.
- Nurdewi. 2022. *Implementasi Personal Branding Smart ASN Perwujudan Bangsa Melayani Di Provinsi Maluku Utara*. Jurnal: SENTRI: Jurnal Riset Ilmiah Vol.1, No. 2 Oktober 2022.

- Panjaitan, Dorothy Rouly dan Ahmad, Aripin. 2017. Buku Ajar Metode Penelitian untuk Bisnis. Bandar Lampung: Lembaga Pengembangan Pembelajaran dan Penjaminan Mutu (LP3M) Universitas Lampung
- Purnomo, Rochmat Aldy. 2016. *Analisis Statistik Ekonomi dan Bisnis dengan SPSS*. Ponorogo: CV Wade Group.
- Rachman, Arif. dkk. 2024. *Metode Penelitian Kuantitatif, Kualitatif & R&D*. Kab. Karawang: CV Saba Jaya Publisher.
- Rahmadi. 2011. Pengantar Metodologi Penelitian. Banjarmasin: Penerbit Antasari Press.
- Rukminingsih, Adnan, Gunawan, dan Latief, Mohammad Adnan. 2020. Metode Penelitian: Penelitian Kuantitatif, Penelitian Kualitatif, Penelitian Tindakan Kelas. Yogyakarta: Erhaka Utama.
- Sahir, Syafrida Hafni. 2021. *Metodologi Penelitian*. Jogyakarta: Penerbit KBM Indonesia.
- Sidiq, Umar dan Choiri, Moh. Miftachul. 2019. *Metode Penelitian Kualitatif di Bidang Pendidikan*. Ponorogo: CV. Nata Karya.
- Warsono, Hardi, Astuti, Retno Sunu, dan Ardiyansah. 2022. *Metode Pengolahan Data Kualitatif Menggunakan Atlas.ti*. Semarang: Program Studi Doktor Administrasi Publik FISIP-UNDIP.
- Widodo, Slamet. dkk. 2023. *Buku Ajar Metode Penelitian*. Pangkal Pinang: Penerbit CV Science Techno Direct.

BAB 4

VISUALISASI DATA DAN PENGAMBILAN KEPUTUSAN

Oleh Lawrence Adi Supriyono

4.1 Pendahuluan

Dalam dunia yang semakin berorientasi pada data, kemampuan untuk mengumpulkan, memahami, dan menyampaikan informasi dari data menjadi lebih penting dari sebelumnya. Salah satu cara utama untuk menjembatani data mentah dengan pemahaman manusia adalah melalui visualisasi data.

Visualisasi data adalah proses mengubah data mentah menjadi representasi grafis seperti grafik, diagram, atau peta, yang memudahkan pemahaman informasi yang kompleks. Dengan visualisasi, data yang besar dan rumit dapat dipahami lebih cepat dan akurat, serta membantu pengguna menemukan pola, hubungan, dan tren yang tersembunyi. Dalam konteks yang lebih modern, visualisasi data bukan hanya tentang estetika, tetapi juga tentang efektivitas dalam menyampaikan informasi yang bisa langsung digunakan untuk pengambilan keputusan.

Pengambilan keputusan berbasis data (data-driven decision making) adalah pendekatan dalam pengambilan keputusan yang mengandalkan analisis data untuk menentukan langkah yang akan diambil, bukan hanya

berdasarkan intuisi atau pengalaman. Visualisasi berperan sangat penting dalam proses ini, karena dengan penyajian yang tepat, data dapat menampilkan tren, pola, anomali, dan peluang yang tersembunyi yang mungkin tidak terlihat melalui angka atau tabel saja.

Menurut Knaflit (2015), visualisasi data yang baik dapat mengubah data yang kompleks menjadi cerita yang mudah dipahami, yang membantu pengambil keputusan untuk lebih cepat merespons situasi yang sedang berlangsung. Hal ini sejalan dengan pemikiran Cairo (2016), yang menekankan bahwa visualisasi yang efektif harus mampu menyampaikan makna yang jelas dan mengurangi kemungkinan distorsi informasi. Sebuah visualisasi yang buruk dapat menyesatkan dan menyebabkan pengambilan keputusan yang salah.

Dalam lingkungan yang semakin digital dan berbasis data, visualisasi data menjadi komponen utama dalam sistem dashboard dan laporan analitik, yang membantu para profesional untuk menyampaikan hasil riset dan analisis kepada berbagai pemangku kepentingan, termasuk pelanggan, investor, dan tim internal. Ware (2012) menyatakan bahwa visualisasi data memungkinkan kita untuk mengubah data besar menjadi insight yang bisa diakses dengan cepat, yang sangat penting dalam pengambilan keputusan yang lebih cepat dan lebih baik.

Dalam konteks bisnis dan sains modern, visualisasi data menjadi komponen kunci dalam sistem dashboard, laporan analitik, dan komunikasi hasil riset (Pratama, 2020). Hal ini tidak hanya membantu dalam pengambilan keputusan internal, tetapi juga dalam menyampaikan hasil kepada pihak eksternal seperti pelanggan, investor, atau pemangku kepentingan.

Adapun karakteristik penting dalam visualisasi data meliputi:

- Kejelasan (*clarity*): Informasi harus mudah dipahami.
- Keakuratan (*accuracy*): Tidak mengubah makna data.
- Efisiensi (*efficiency*): Menyajikan data sebanyak mungkin dalam ruang visual yang terbatas.
- Daya tarik visual (*aesthetic appeal*): Memotivasi audiens untuk memperhatikan informasi yang disajikan.

Dengan kemajuan teknologi, berbagai tools visualisasi data kini tersedia seperti Tableau, Power BI, hingga Python dengan pustaka seperti Matplotlib dan Seaborn. Setiap alat ini menawarkan pendekatan yang berbeda untuk membantu memvisualisasikan data dari berbagai sumber dan dalam berbagai skala.

Secara keseluruhan, visualisasi data adalah seni dan ilmu yang menghubungkan dunia data dengan manusia, mengubah kumpulan angka menjadi cerita visual yang bermakna dan mendalam (Kirk, 2016).

4.2 Tujuan Visualisasi Data

Tujuan utama dari visualisasi data adalah menyampaikan informasi secara efektif melalui media visual agar dapat dipahami dengan cepat dan akurat oleh manusia. Ketika data hanya disajikan dalam bentuk tabel atau teks, proses analisis menjadi lebih berat dan memerlukan waktu yang lebih lama. Dengan menggunakan visualisasi, kompleksitas data dapat direduksi menjadi representasi yang lebih sederhana dan intuitif.

Menurut Gani dan Yusuf (2019), ada beberapa tujuan utama dalam melakukan visualisasi data, yaitu:

1. Mempermudah Pemahaman Data Visualisasi mengubah angka-angka menjadi bentuk visual seperti grafik batang, diagram garis, atau peta panas, yang membuat pola dan hubungan dalam data lebih cepat dikenali.
2. Meningkatkan Kemampuan Pengambilan Keputusan Dengan menyajikan informasi kunci secara jelas dan ringkas, visualisasi data membantu pengambil keputusan untuk mengevaluasi situasi dengan lebih baik dan memilih langkah yang tepat (Kirk, 2016).
3. Mengungkap Pola dan Tren Tersembunyi Banyak pola atau anomali dalam dataset besar yang tidak terlihat melalui analisis tradisional dapat dengan cepat terdeteksi melalui visualisasi (Pratama, 2020).
4. Komunikasi Data kepada Audiens Non-Teknis Visualisasi menjadi jembatan antara data teknis dan audiens awam. Penyampaian data yang rumit dapat dibuat lebih mudah dicerna dan menarik untuk berbagai pihak, termasuk manajemen, investor, atau pelanggan (Darmawan, 2018).
5. Meningkatkan Interaktivitas dalam Eksplorasi Data Dengan perkembangan teknologi, visualisasi data kini tidak hanya bersifat statis, melainkan juga interaktif. Tools seperti Tableau dan Power BI memungkinkan pengguna untuk berinteraksi langsung dengan data, mengeksplorasi detail lebih lanjut sesuai kebutuhan analisis (Ware, 2012).

Visualisasi data yang dirancang dengan baik mampu menciptakan "storytelling" data yang kuat, di mana setiap visualisasi menceritakan suatu narasi yang mendorong pemahaman yang lebih dalam terhadap konteks dan makna data tersebut (Knaflic, 2015). Seiring dengan semakin pentingnya data dalam pengambilan keputusan strategis, kemampuan dalam membuat visualisasi yang efektif menjadi keterampilan esensial dalam dunia profesional saat ini.

4.3 Prinsip-Prinsip Visualisasi Data yang Efektif

Membuat visualisasi data yang efektif bukan sekadar memilih grafik yang menarik. Sebaliknya, visualisasi harus didesain dengan prinsip-prinsip tertentu agar mampu menyampaikan informasi dengan jelas, akurat, dan bermakna (Evergreen, 2017). Visualisasi yang baik harus memperhatikan baik aspek estetika maupun aspek fungsionalitas data.

Berikut beberapa prinsip penting dalam menciptakan visualisasi data yang efektif:

4.3.1 Kejelasan (*Clarity*)

Visualisasi harus mampu mengkomunikasikan informasi secara sederhana tanpa membingungkan audiens. Kesederhanaan desain membantu pembaca fokus pada inti pesan, bukan pada elemen dekoratif yang tidak perlu (Knaflic, 2015).

4.3.2 Akurasi (*Accuracy*)

Data harus disajikan secara tepat, tidak boleh menyesatkan atau memanipulasi persepsi. Misalnya, penggunaan skala grafik harus konsisten agar tidak membesar-besarkan atau mengecilkan suatu perbedaan.

4.3.3 Fokus pada Pesan Utama

Setiap elemen dalam visualisasi harus mengarahkan perhatian *audiens* pada pesan inti. Penggunaan warna, ukuran font, dan layout harus memperjelas narasi data yang ingin disampaikan (Cairo, 2016).

4.3.4 Meminimalkan “*Chartjunk*”

“*Chartjunk*” adalah istilah untuk elemen visual yang tidak perlu seperti latar belakang berlebihan, gambar dekoratif, atau efek 3D yang mengaburkan data. Visualisasi yang bersih akan mempercepat pemahaman.

4.3.5 Memilih Jenis Grafik yang Tepat

Jenis visualisasi harus dipilih berdasarkan karakteristik data dan tujuan analisis. Misalnya, data time series lebih tepat divisualisasikan dengan grafik garis, sementara proporsi kategori lebih sesuai dengan diagram batang atau pie chart (Kirk, 2016).

4.3.6 Konsistensi Desain

Penggunaan warna, bentuk, ukuran teks, dan simbol harus konsisten di seluruh visualisasi agar audiens mudah menghubungkan informasi yang terkait (Evergreen, 2017).

4.3.7 Mendorong Interaktivitas

Dalam konteks modern, visualisasi tidak hanya statis, tetapi juga bisa bersifat interaktif. Ini memungkinkan audiens mengeksplorasi data lebih dalam berdasarkan kebutuhan masing-masing, terutama dalam dashboard analitik (Ware, 2012).

Visualisasi data yang efektif mempertemukan fungsi dan estetika untuk menciptakan komunikasi yang kuat. Mengabaikan prinsip-prinsip ini dapat membuat visualisasi menjadi tidak berguna, bahkan menyesatkan audiens.

4.4 Alat-Alat untuk Visualisasi Data

Dalam dunia analitik modern, memilih teknik visualisasi yang tepat sangat penting untuk memastikan data dapat diinterpretasikan secara optimal. Setiap jenis data dan tujuan analisis membutuhkan pendekatan visual yang berbeda (Yau, 2013). Oleh karena itu, memahami berbagai teknik visualisasi menjadi modal utama dalam menyampaikan insight secara efektif.

Berikut beberapa teknik visualisasi data yang umum digunakan:

4.4.1 Grafik Garis (*Line Chart*)

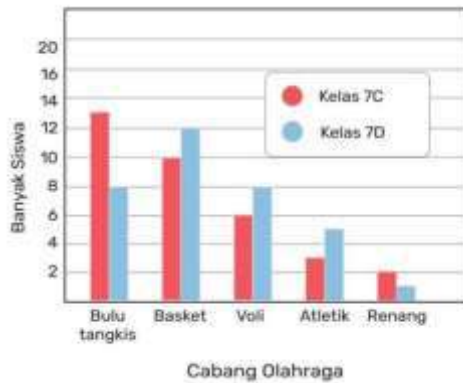


Gambar 4. 1 Contoh Grafik Garis

Sumber : [datavizproject-com](https://datavizproject.com)

Grafik garis digunakan untuk menunjukkan perubahan data dari waktu ke waktu (*trend analysis*). Grafik ini efektif menampilkan hubungan antar variabel dan fluktuasi data sepanjang periode tertentu. Sebagai contoh, perubahan jumlah pengguna aplikasi bulanan dapat divisualisasikan dengan grafik garis.

4.4.2 Diagram Batang (*Bar Chart*)

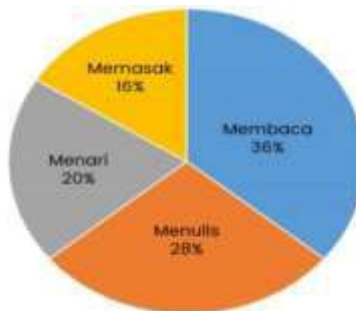


Gambar 4. 2 Contoh Grafik Batang

Sumber : detik.com

Diagram batang adalah pilihan ideal untuk membandingkan data antar kategori. Dengan batang vertikal atau horizontal, perbandingan antar nilai menjadi lebih jelas. Teknik ini banyak digunakan dalam laporan bisnis, survei, dan analisis pasar (Gani dan Yusuf, 2019).

4.4.3. Diagram Lingkaran (*Pie Chart*)

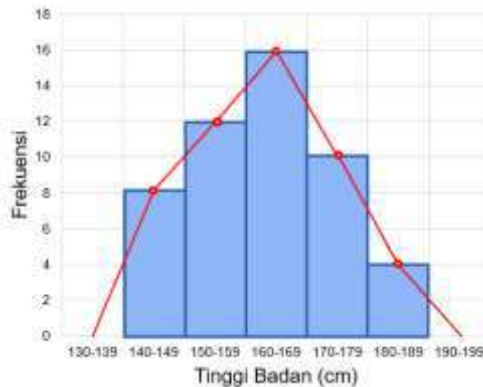


Gambar 4. 3 Contoh Grafik Diagram Lingkaran

Sumber : katadata.co.id

Diagram lingkaran digunakan untuk menunjukkan proporsi atau persentase dari suatu keseluruhan. Meskipun populer, penggunaan pie chart harus dibatasi untuk jumlah kategori yang sedikit agar tetap mudah dibaca (Knafllic, 2015).

4.4.4 Histogram

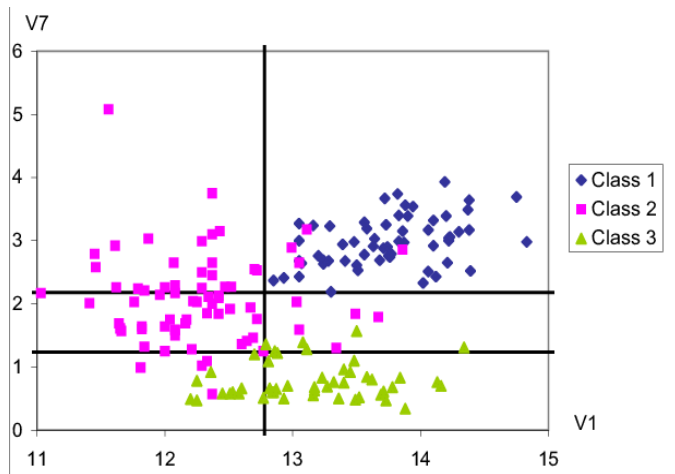


Gambar 4. 4 Contoh Grafik Histogram

Sumber : roboguru.ruangguru.com

Histogram digunakan untuk menunjukkan distribusi frekuensi data numerik yang dikelompokkan dalam interval tertentu. Teknik ini sangat berguna untuk memahami pola sebaran data, mendeteksi outlier, atau menilai normalitas suatu dataset (Pratama, 2020).

4.4.5 Diagram Pencar (*Scatter Plot*)

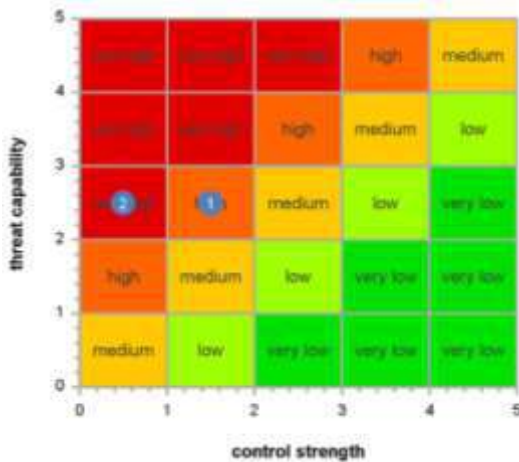


Gambar 4.5 Contoh Grafik Scatter Plot

Sumber : roboguru.ruangguru.com

Scatter plot efektif untuk mengidentifikasi hubungan antara dua variabel numerik. Pola yang muncul, seperti korelasi positif atau negatif, dapat diinterpretasikan dengan lebih cepat melalui teknik ini (Ware, 2012).

4.4.6 Heatmap

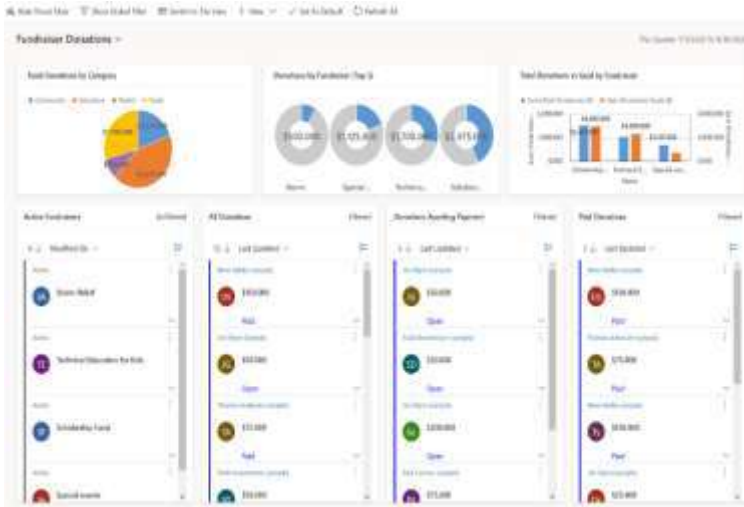


Gambar 4. 5 Contoh Grafik *Heatmap*

Sumber : bizzdesign.com

Heatmap menyajikan data dalam bentuk warna, di mana intensitas warna menunjukkan nilai data. Ini berguna untuk menganalisis hubungan kompleks antar variabel, seperti korelasi dalam dataset besar (Wickham dan Grolemond, 2017).

4.4.7 Dashboard Interaktif

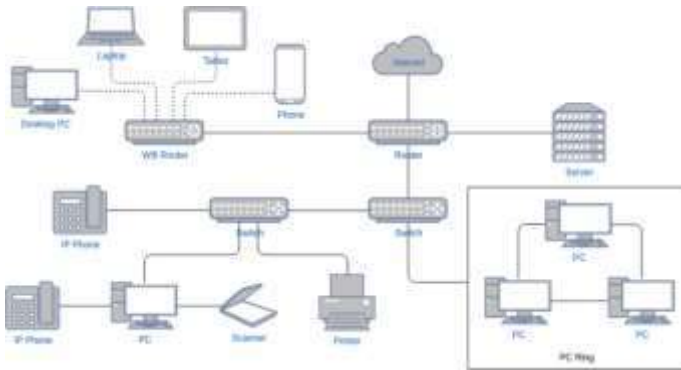


Gambar 4. 6 Contoh Dashboard Interaktif

Sumber : learn.microsoft.com

Dashboard menggabungkan berbagai teknik visualisasi dalam satu tampilan terpadu, memungkinkan pengguna untuk memantau dan menganalisis berbagai metrik secara real-time. *Platform* seperti Tableau, Power BI, dan Google Data Studio memanfaatkan teknik ini untuk mendukung pengambilan keputusan berbasis data (Sulaiman, 2021).

4.4.8 Diagram Jaringan (*Network Diagram*)



Gambar 4. 7 Contoh Diagram Jaringan

Sumber : online.visual-paradigm.com

Untuk merepresentasikan hubungan antar entitas (seperti hubungan sosial atau koneksi antar perangkat dalam jaringan komputer), diagram jaringan menjadi sangat efektif. Node dan edge digunakan untuk menggambarkan entitas dan hubungan di antara mereka (Cairo, 2016).

Pemilihan teknik visualisasi yang tepat harus mempertimbangkan karakteristik data, tujuan analisis, serta siapa audiensnya. Kesalahan dalam memilih metode visualisasi dapat menyebabkan interpretasi yang keliru dan pengambilan keputusan yang salah (Evergreen, 2017).

4.5 Teknik Visualisasi Data

Seiring dengan berkembangnya teknologi, berbagai alat atau tools visualisasi data telah muncul untuk mempermudah analisis dan penyajian data. Alat-alat ini membantu analis, ilmuwan data, dan profesional lainnya dalam menggali insight dari data dan menyajikannya dengan cara yang mudah dipahami oleh audiens. Pemilihan tools yang tepat bergantung pada kebutuhan, kompleksitas data, dan tingkat interaktivitas yang diinginkan.

Berikut adalah beberapa tools visualisasi data populer yang banyak digunakan saat ini:

4.5.1 *Tableau*

Tableau adalah salah satu alat visualisasi data paling populer di dunia. Dengan antarmuka yang intuitif, pengguna dapat menghubungkan *Tableau* langsung dengan berbagai sumber data dan membuat visualisasi interaktif secara mudah. *Tableau* menyediakan berbagai jenis grafik, termasuk peta geospasial, grafik garis, bar chart, dan scatter plot. Alat ini sering digunakan oleh perusahaan besar untuk dashboard analitik real-time (Sulaiman, 2021).

4.5.2 *Power BI*

Power BI adalah alat visualisasi data dari Microsoft yang terintegrasi dengan berbagai produk Microsoft lainnya seperti Excel dan Azure. Dengan *Power BI*, pengguna dapat membuat laporan dan dashboard interaktif yang dapat diakses secara online. Salah satu keunggulan *Power BI* adalah kemampuannya untuk bekerja dengan data yang besar dan kompleks, serta menyediakan analisis berbasis AI (Fisher, 2018).

4.5.3 Google Data Studio

Google Data Studio adalah platform visualisasi data berbasis cloud yang memungkinkan pengguna untuk membuat laporan dan dashboard yang dapat dibagikan secara online. Alat ini terintegrasi dengan berbagai produk Google seperti Google Analytics dan Google Sheets, sehingga sangat berguna untuk analisis web dan pemasaran digital. Google Data Studio menyediakan berbagai template dan alat kustomisasi untuk memenuhi kebutuhan analisis yang beragam (Evergreen, 2017).

4.5.4 QlikView

QlikView adalah alat visualisasi data yang menawarkan analitik berbasis data secara interaktif. Dengan menggunakan teknologi associative data model, QlikView memungkinkan pengguna untuk mengeksplorasi data dari berbagai perspektif dan menemukan pola yang tersembunyi. Alat ini banyak digunakan dalam bisnis untuk analisis data yang lebih mendalam (Yau, 2013).

4.5.5 R dan ggplot2

R adalah bahasa pemrograman yang sering digunakan untuk analisis statistik, namun juga memiliki kemampuan visualisasi data yang sangat kuat melalui package ggplot2. Dengan ggplot2, pengguna dapat membuat berbagai jenis grafik dan plot yang sangat customizable. R dengan ggplot2 sering digunakan oleh data scientist dan peneliti untuk visualisasi data yang kompleks dan analisis statistik tingkat lanjut (Wickham, 2016).

4.5.6. D3.js

D3.js adalah pustaka JavaScript yang memungkinkan pembuatan visualisasi data interaktif berbasis web. D3.js memberi pengguna kontrol penuh terhadap setiap elemen visualisasi, sehingga memungkinkan desain yang sangat dinamis dan interaktif. Alat ini sangat berguna untuk pengembang yang ingin membuat visualisasi data yang sangat custom dan interaktif di situs web atau aplikasi (Bostock et al., 2011).

4.5.7 Plotly

Plotly adalah platform visualisasi data yang dapat digunakan untuk membuat grafik interaktif dan publikasi ilmiah. *Platform* ini mendukung berbagai bahasa pemrograman, termasuk Python dan R. *Plotly* dikenal karena kemampuannya dalam membuat visualisasi yang interaktif dan mudah diintegrasikan dengan aplikasi berbasis web (Pascariu et al., 2019).

4.5.8 Excel

Meskipun bukan alat visualisasi data khusus, Excel tetap menjadi salah satu alat yang paling sering digunakan untuk membuat visualisasi data sederhana, seperti grafik batang, grafik garis, dan pie chart. Excel menawarkan kemudahan bagi pemula untuk memvisualisasikan data tanpa memerlukan perangkat lunak khusus, namun fungsinya terbatas dibandingkan dengan alat yang lebih canggih seperti Tableau atau Power BI (Gani dan Yusuf, 2019).

Pemilihan alat visualisasi data yang tepat tergantung pada kebutuhan spesifik pengguna dan jenis data yang dikelola. Tools seperti Tableau dan Power BI lebih cocok untuk analisis berbasis bisnis dengan kebutuhan dashboard interaktif, sementara alat seperti R dan ggplot2 cocok bagi pengguna yang membutuhkan analisis statistik mendalam. Sementara itu, D3.js menawarkan kontrol penuh atas desain visual, membuatnya ideal untuk pengembang yang ingin membuat visualisasi data kustom berbasis web.

4.6 Tantangan dalam Visualisasi Data

Visualisasi data menawarkan berbagai manfaat dalam mempermudah pemahaman informasi, namun proses ini juga memiliki tantangan tersendiri. Tantangan dalam visualisasi data dapat berasal dari berbagai aspek, mulai dari pengumpulan data yang akurat hingga penyampaian pesan yang jelas kepada audiens. Memahami tantangan-tantangan ini adalah langkah pertama untuk menciptakan visualisasi data yang efektif dan berimpact.

4.6.1 Memilih Teknik Visualisasi yang Tepat

Salah satu tantangan terbesar dalam visualisasi data adalah memilih teknik yang paling sesuai untuk jenis data yang ada. Tidak semua grafik atau diagram cocok untuk setiap tipe data. Misalnya, grafik garis lebih baik untuk menunjukkan tren data dalam jangka waktu tertentu, sementara diagram batang lebih efektif untuk perbandingan antar kategori. Kesalahan dalam memilih teknik visualisasi dapat mengarah pada interpretasi yang keliru dan menyulitkan audiens dalam memahami informasi.

4.6.2 Menghindari *Overload* Informasi

Ketika menyajikan data dalam bentuk visual, sangat mudah untuk terjebak dalam godaan untuk memasukkan terlalu banyak informasi dalam satu visualisasi. *Overload* informasi (*information overload*) dapat mengaburkan pesan utama yang ingin disampaikan. Visualisasi yang terlalu rumit atau padat dengan informasi justru dapat menyebabkan kebingungan dan mengurangi kejelasan (Knafllic, 2015). Oleh karena itu, penting untuk menyederhanakan visualisasi dengan hanya menampilkan data yang relevan dan mudah dipahami.

4.6.3 Pengaruh Skala dan Sumbu

Skala dan sumbu dalam visualisasi dapat sangat memengaruhi cara data dipersepsikan. Misalnya, menggunakan skala logaritmik untuk data yang bersifat eksponensial dapat membantu menampilkan data dengan rentang nilai yang sangat besar secara lebih jelas. Namun, jika salah menggunakan skala atau sumbu, visualisasi bisa menipu audiens dan menyebabkan interpretasi yang salah (Evergreen, 2017). Oleh karena itu, pemilihan skala dan sumbu yang sesuai sangat penting untuk menjaga integritas visualisasi.

4.6.4 Ketergantungan pada Kualitas Data

Visualisasi data sangat bergantung pada kualitas data yang digunakan. Data yang buruk atau tidak lengkap dapat menghasilkan visualisasi yang menyesatkan dan merugikan proses pengambilan keputusan. Data yang salah, tidak akurat, atau tidak diperbarui dengan baik dapat mengaburkan wawasan yang seharusnya diperoleh dari visualisasi tersebut. Oleh karena itu, penting untuk memastikan kualitas data yang digunakan dalam proses visualisasi (Yau, 2013).

4.6.5 Beragam Audiens dengan Latar Belakang yang Berbeda

Visualisasi data sering kali digunakan oleh berbagai audiens, mulai dari eksekutif perusahaan hingga peneliti teknis. Masing-masing kelompok audiens ini memiliki tingkat pemahaman yang berbeda tentang data dan cara visualisasi yang digunakan. Oleh karena itu, tantangan besar dalam visualisasi data adalah menyampaikan informasi dengan cara yang dapat dipahami oleh berbagai audiens. Hal ini sering kali melibatkan penyederhanaan data tanpa mengurangi substansi yang penting (Cairo, 2016).

4.6.6 Menjaga Interaktivitas Tanpa Mengorbankan Kejelasan

Dalam visualisasi interaktif, pengguna dapat berinteraksi dengan data melalui filter, zoom, dan klik untuk mengeksplorasi lebih dalam. Meskipun interaktivitas meningkatkan keterlibatan, ini juga dapat menjadi tantangan jika tidak dirancang dengan baik. Terlalu banyak opsi interaktif dapat membingungkan pengguna dan mengurangi kejelasan visualisasi. Menemukan keseimbangan yang tepat antara interaktivitas dan kejelasan adalah hal yang penting dalam membuat visualisasi yang efektif (Pascariu et al., 2019).

4.6.7 Menerjemahkan Data ke dalam Bentuk Visual yang Akrab

Tidak semua data mudah diterjemahkan menjadi visual yang akrab. Beberapa jenis data, seperti data non-numerik atau data yang sangat kompleks, memerlukan representasi visual yang lebih kreatif. Menerjemahkan data yang kompleks ke dalam visualisasi yang sederhana namun akurat memerlukan keterampilan desain yang

baik serta pemahaman yang mendalam tentang data tersebut. Ini adalah tantangan utama bagi desainer visualisasi data (Pratama, 2020).

Tantangan-tantangan ini menuntut para profesional visualisasi data untuk tidak hanya memiliki keterampilan teknis dalam menggunakan alat dan perangkat lunak, tetapi juga untuk memiliki pemahaman yang mendalam tentang data yang mereka visualisasikan dan audiens yang mereka tuju. Mengatasi tantangan-tantangan ini adalah langkah krusial dalam menciptakan visualisasi data yang efektif dan dapat digunakan untuk mendukung pengambilan keputusan yang tepat.

4.7 Manfaat Visualisasi Data dalam Pengambilan Keputusan

Visualisasi data bukan hanya alat untuk menampilkan informasi, tetapi juga memiliki peran yang sangat penting dalam mendukung proses pengambilan keputusan. Dalam dunia yang semakin dipenuhi dengan volume data yang besar, kemampuan untuk menyajikan data dalam bentuk yang mudah dipahami dan dapat diinterpretasikan dengan cepat sangat diperlukan. Visualisasi data dapat membantu individu dan organisasi untuk mengambil keputusan yang lebih baik, lebih cepat, dan lebih informasional.

4.7.1 Mempermudah Pemahaman Data yang Kompleks

Salah satu manfaat utama visualisasi data adalah kemampuannya untuk menyederhanakan data yang kompleks dan besar menjadi format yang lebih mudah dipahami. Data mentah yang besar dan tidak terstruktur seringkali sulit untuk diinterpretasikan tanpa bantuan visualisasi. Melalui penggunaan grafik, peta, dan

diagram interaktif, visualisasi data memungkinkan audiens untuk mengidentifikasi pola, tren, dan anomali yang mungkin sulit ditemukan dalam data tabular (Cairo, 2016). Sebagai contoh, grafik garis yang menunjukkan tren penjualan tahunan dapat dengan mudah mengungkapkan pola musiman atau fluktuasi yang signifikan yang mungkin tidak tampak jika hanya dilihat dalam format tabel.

4.7.2 Mempercepat Pengambilan Keputusan

Dalam konteks bisnis dan organisasi, waktu adalah faktor yang sangat penting. Visualisasi data memungkinkan pengambilan keputusan yang lebih cepat dengan mengurangi waktu yang diperlukan untuk menganalisis data secara manual. Pengambilan keputusan yang cepat sangat penting dalam banyak situasi, terutama dalam pasar yang kompetitif atau saat merespons perubahan yang cepat. Dengan menampilkan data dalam bentuk yang jelas dan mudah dicerna, manajer atau pengambil keputusan dapat dengan segera melihat gambaran besar dan membuat keputusan yang lebih cepat dan tepat.

4.7.3 Meningkatkan Keakuratan dan Validitas Keputusan

Visualisasi data juga dapat meningkatkan keakuratan keputusan yang diambil dengan memberikan wawasan yang lebih jelas dan lebih transparan. Ketika data dipresentasikan dengan cara yang benar dan konsisten, pengguna dapat lebih mudah mengidentifikasi faktor-faktor yang memengaruhi hasil dan membuat keputusan yang didasarkan pada pemahaman yang lebih baik tentang situasi yang ada. Misalnya, analisis data penjualan yang divisualisasikan dapat membantu manajer dalam menentukan apakah sebuah produk memerlukan promosi tambahan atau

jika tren penurunan penjualan berkaitan dengan faktor eksternal seperti perubahan ekonomi (Yau, 2013).

4.7.4 Menunjang Kolaborasi yang Lebih Baik

Visualisasi data tidak hanya bermanfaat untuk pengambil keputusan tunggal, tetapi juga sangat mendukung kolaborasi dalam tim. Dalam banyak organisasi, keputusan penting sering melibatkan berbagai pihak yang memiliki latar belakang dan keahlian yang berbeda. Visualisasi data yang jelas dapat menjadi alat yang efektif untuk menyampaikan informasi secara kolektif kepada berbagai pihak, memungkinkan diskusi yang lebih produktif dan kolaborasi yang lebih baik. Semua anggota tim dapat melihat dan menganalisis data secara bersama-sama, membantu mereka untuk merumuskan solusi yang lebih holistik (Knafllic, 2015).

4.7.5 Membantu Memprediksi Tren Masa Depan

Salah satu keunggulan visualisasi data adalah kemampuannya untuk memprediksi tren masa depan berdasarkan data historis. Dengan menggunakan teknik visualisasi yang tepat, seperti grafik tren atau peta panas, organisasi dapat melihat proyeksi yang didasarkan pada data masa lalu dan mempersiapkan diri untuk menghadapi tantangan atau peluang di masa depan. Dalam konteks bisnis, ini dapat mencakup proyeksi penjualan, analisis biaya, atau perkiraan permintaan pasar. Ini sangat membantu dalam perencanaan strategis dan pengambilan keputusan jangka panjang (Pascariu et al., 2019).

4.7.6 Meningkatkan Aksesibilitas dan Transparansi

Visualisasi data juga berperan penting dalam meningkatkan aksesibilitas informasi. Dengan menggunakan alat visualisasi yang baik, data yang kompleks dapat disajikan secara jelas dan dapat diakses oleh berbagai audiens, termasuk mereka yang tidak memiliki latar belakang teknis. Hal ini membantu dalam membuat informasi lebih transparan dan mudah dimengerti oleh semua pemangku kepentingan yang terlibat dalam pengambilan keputusan (Pratama, 2020).

4.7.7 Mendukung Keputusan Berdasarkan Data (*Data-Driven Decision Making*)

Salah satu konsep penting dalam dunia bisnis modern adalah pengambilan keputusan berbasis data (*data-driven decision making*). Visualisasi data memainkan peran kunci dalam mendorong keputusan berbasis data dengan menyediakan gambaran yang jelas tentang informasi yang relevan. Dengan visualisasi yang tepat, pengambil keputusan dapat mengurangi bias pribadi dan intuisi, yang seringkali dapat memengaruhi proses pengambilan keputusan secara negatif. Keputusan yang diambil dengan dasar data cenderung lebih objektif dan dapat dipertanggungjawabkan (Evergreen, 2017).

Manfaat visualisasi data sangat besar dalam meningkatkan kualitas dan kecepatan pengambilan keputusan. Dengan kemampuannya untuk menyederhanakan data kompleks, mempercepat proses analisis, dan memberikan wawasan yang lebih dalam, visualisasi data menjadi alat yang sangat penting dalam mendukung pengambilan keputusan yang lebih baik dan lebih tepat dalam berbagai konteks.

4.8 Tantangan dalam Visualisasi Data

Meskipun visualisasi data memiliki banyak manfaat dalam pengambilan keputusan, proses ini tidak bebas dari tantangan. Berbagai kendala dapat muncul, mulai dari pemilihan visualisasi yang tidak tepat hingga keterbatasan teknologi atau literasi data pengguna. Oleh karena itu, memahami tantangan-tantangan ini menjadi penting agar visualisasi data yang disajikan tidak menyesatkan dan tetap efektif.

4.8.1 Pemilihan Visualisasi yang Tidak Sesuai

Salah satu tantangan utama adalah memilih jenis visualisasi yang tidak sesuai dengan data yang ditampilkan. Misalnya, penggunaan pie chart untuk data dengan banyak kategori bisa menyebabkan kesulitan dalam membedakan nilai antar kategori. Kesalahan pemilihan grafik dapat mengarah pada interpretasi yang salah atau membingungkan pengguna (Kosara, 2016).

4.8.2 Kualitas dan Kebersihan Data (*Data Quality*)

Visualisasi yang baik sangat bergantung pada kualitas data. Data yang tidak lengkap, mengandung outlier yang tidak ditangani, atau tidak dibersihkan dengan baik dapat menyebabkan visualisasi yang menyesatkan. Tantangan ini sering kali terjadi karena pengumpulan data yang tidak konsisten atau kurangnya proses validasi sebelum visualisasi dilakukan (Kimball & Ross, 2013).

4.8.3 Overload Informasi (*Information Overload*)

Terlalu banyak informasi dalam satu tampilan visual bisa menjadi kontraproduktif. Alih-alih memperjelas data, visualisasi yang padat dan penuh bisa menyulitkan pengguna dalam memahami insight utama.

Ini biasanya disebabkan oleh keinginan untuk menampilkan “semua hal” dalam satu layar tanpa menyaring atau mengorganisasi data secara efektif.

4.8.4 Kurangnya Literasi Data Pengguna

Tidak semua pengguna memiliki tingkat pemahaman data yang sama. Tantangan ini muncul terutama dalam organisasi yang memiliki pemangku kepentingan dari latar belakang non-teknis. Visualisasi yang tampak jelas bagi analis data mungkin tidak dipahami oleh manajer atau stakeholder yang tidak terbiasa dengan istilah teknis atau tipe grafik tertentu (Boy et al., 2015).

4.8.5 Keterbatasan Alat Visualisasi

Terkadang, alat visualisasi data yang digunakan tidak memiliki fitur yang cukup fleksibel untuk mengakomodasi jenis data tertentu atau interaktivitas yang dibutuhkan. Hal ini bisa membatasi kemampuan pengguna dalam menjelajah data secara mendalam. Beberapa perangkat lunak memiliki antarmuka yang kompleks, kurang user-friendly, atau tidak mendukung integrasi data dari berbagai sumber (Kirk, 2019).

4.8.6 Bias dan Manipulasi Data

Visualisasi yang disengaja dibuat untuk memanipulasi persepsi atau membingungkan pengguna juga merupakan tantangan yang cukup serius. Ini dapat terjadi karena kesalahan desain visualisasi atau bahkan sengaja dilakukan untuk menyembunyikan atau memutarbalikkan fakta. Praktik semacam ini menurunkan kredibilitas data dan keputusan yang dihasilkan (Pandey et al., 2015).

4.9 Best Practices: Membuat Visualisasi yang Memberikan Dampak

Untuk memahami secara nyata dampak visualisasi data dalam pengambilan keputusan, bagian ini menyajikan beberapa studi kasus dari berbagai industri yang telah berhasil menerapkan visualisasi data sebagai alat strategis. Studi kasus ini menunjukkan bagaimana representasi visual dari data dapat mengungkap pola tersembunyi, memprediksi tren, dan mempercepat pengambilan keputusan.

4.9.1 Sektor Kesehatan: Pelacakan Penyebaran Penyakit

Pada masa pandemi COVID-19, *Johns Hopkins University* mengembangkan dashboard visualisasi interaktif yang menampilkan penyebaran kasus secara global secara real-time. Peta interaktif dan grafik tren harian ini tidak hanya membantu masyarakat umum memahami kondisi terkini, tetapi juga menjadi sumber penting bagi pemerintah dan lembaga kesehatan dalam mengambil tindakan mitigasi (Dong, Du dan Gardner, 2020). Visualisasi ini mempermudah identifikasi daerah rawan dan efektivitas kebijakan karantina wilayah.

4.9.2 Ritel dan *E-commerce*: Prediksi Penjualan

Amazon dan berbagai perusahaan *e-commerce* lainnya menggunakan visualisasi data untuk memahami perilaku pelanggan. Dengan memanfaatkan heatmap, grafik tren penjualan, dan segmentasi produk berdasarkan lokasi, mereka dapat mengoptimalkan stok dan meningkatkan rekomendasi produk. Contohnya, Amazon memvisualisasikan pergerakan produk dalam rantai pasok menggunakan grafik berbasis aliran (flow charts) untuk mendeteksi potensi keterlambatan dan melakukan penyesuaian logistik secara proaktif (Marr, 2017).

4.9.3 Pemerintahan: Transparansi dan Kinerja Layanan Publik

Pemerintah Kota Jakarta melalui platform *Jakarta Smart City* memanfaatkan dashboard visualisasi untuk menampilkan data kepadatan lalu lintas, kebersihan, banjir, dan aduan masyarakat secara interaktif. Peta digital dan grafik tren aduan warga per kecamatan membantu pengambilan keputusan berbasis data oleh dinas terkait. Ini menciptakan transparansi publik dan memungkinkan partisipasi warga secara lebih aktif (Pemerintah Provinsi DKI Jakarta, 2021).

4.9.4 Bidang Keuangan: Manajemen Risiko

Perusahaan keuangan dan perbankan seperti Bank Mandiri dan BCA menggunakan dashboard visualisasi untuk memantau performa portofolio, risiko kredit, dan tren transaksi nasabah. Visualisasi dalam bentuk scatter plot dan matriks korelasi membantu analis mendeteksi anomali dalam transaksi atau potensi fraud lebih cepat dibandingkan jika data hanya disajikan dalam bentuk tabel numerik (Chen et al., 2012).

4.10 Kesimpulan

Visualisasi data merupakan komponen penting dalam proses analitik modern karena menjembatani antara kompleksitas data dan pemahaman manusia. Sepanjang bab ini, telah dijelaskan bahwa visualisasi data bukan hanya soal mempercantik tampilan, tetapi berperan vital dalam membantu individu dan organisasi membuat keputusan yang lebih cepat, tepat, dan berbasis bukti.

Proses visualisasi dimulai dari pemahaman jenis data, pemilihan grafik yang sesuai, hingga penggunaan alat bantu visualisasi seperti Tableau, Power BI, dan Python dengan

pustaka Matplotlib dan Seaborn. Penggunaan prinsip desain visual yang baik seperti pemilihan warna, penggunaan hierarki visual, dan keterbacaan sangat berperan dalam menyampaikan informasi dengan efektif.

Visualisasi yang baik dapat mengungkap pola tersembunyi, mempermudah komunikasi antar pemangku kepentingan, hingga mendorong tindakan strategis. Namun demikian, berbagai tantangan seperti salah pilih grafik, literasi data yang rendah, hingga potensi manipulasi visual tetap perlu diantisipasi. Studi kasus dari sektor kesehatan, pemerintahan, keuangan, hingga e-commerce menunjukkan bahwa visualisasi data telah menjadi bagian integral dari proses pengambilan keputusan di dunia nyata. Setiap keberhasilan dalam penggunaan visualisasi tidak hanya bergantung pada teknologi yang digunakan, tetapi juga pada pemahaman konteks, kualitas data, dan integritas pengguna dalam menyajikan informasi secara objektif. Dengan menguasai visualisasi data secara menyeluruh—baik dari sisi teknis maupun etis—seorang analis data tidak hanya mampu menyajikan informasi, tetapi juga memengaruhi arah strategi dan kebijakan dalam organisasi secara positif dan berkelanjutan.

DAFTAR PUSTAKA

- Aldila, A.S. and Supriyono, L.A., 2024. *Predictive modeling of COVID-19 spread with machine learning: A focus on decision tree accuracy*. Jurnal Teknik Informatika Unika Santo Thomas.
- Boy, J., Rensink, R. A., Bertini, E. and Fekete, J. D. (2015) 'A Principled Way of Assessing Visualization Literacy', IEEE Transactions on Visualization and Computer Graphics, 20(12), pp. 1963–1972.
- Cairo, A. (2016) *The Truthful Art: Data, Charts, and Maps for Communication*. Berkeley, CA: New Riders.
- Chen, M., Ebert, D., Hagen, H., Laramée, R., Van Liere, R., Ma, K.-L., Ribarsky, W., Scheuermann, G. and Silver, D. (2012) 'Data, Information, and Knowledge in Visualization', IEEE Computer Graphics and Applications, 29(1), pp. 12–19.
- Darmawan, D. (2018) *Teknologi Pembelajaran: Teori dan Praktik*. Bandung: Remaja Rosdakarya.
- Dong, E., Du, H. and Gardner, L. (2020) 'An interactive web-based dashboard to track COVID-19 in real time', The Lancet Infectious Diseases, 20(5), pp. 533–534.
- Evergreen, S. D. H. (2017) *Effective Data Visualization: The Right Chart for the Right Data*. Thousand Oaks, CA: Sage Publications.
- Fisher, M. (2018) *Power BI for the Excel Analyst: Real-World Data Visualization*. Hoboken, NJ: Wiley.

- Gani, A. and Yusuf, E. S. (2019) Analisis Data dan Visualisasi dengan Python. Jakarta: Elex Media Komputindo.
- Kimball, R. and Ross, M. (2013) The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. 3rd edn. Indianapolis, IN: Wiley.
- Kirk, A. (2016) Data Visualisation: A Handbook for Data Driven Design. London: Sage Publications.
- Kosara, R. (2016) 'Presentation-Oriented Visualization Techniques', IEEE Computer Graphics and Applications, 36(1), pp. 80–85.
- Knafllic, C. N. (2015) Storytelling with Data: A Data Visualization Guide for Business Professionals. Hoboken, NJ: Wiley.
- Marr, B. (2017) Data Strategy: How to Profit from a World of Big Data, Analytics and the Internet of Things. London: Kogan Page.
- Pandey, A. V., Manivannan, A., Nov, O., Satterthwaite, M. and Bertini, E. (2015) 'The Persuasive Power of Data Visualization', IEEE Transactions on Visualization and Computer Graphics, 20(12), pp. 2211–2220.
- Pascariu, G. et al. (2019) 'Plotly for Data Science: A Practical Approach', Journal of Data Science and Analytics, 12(3), pp. 45-62.
- Pemerintah Provinsi DKI Jakarta. (2021) 'Dashboard Jakarta Smart City'. Tersedia di: <https://smartcity.jakarta.go.id> [Diakses 29 April 2025].

- Pratama, R. A. (2020) Data Science dengan Python: Analisis dan Visualisasi Data untuk Pemula. Jakarta: Penerbit Andi.
- Sulaiman, A. (2021) Data Analytics dengan Power BI: Panduan Praktis untuk Pemula. Yogyakarta: Penerbit Andi.
- Ware, C. (2012) Information Visualization: Perception for Design. 3rd ed. Burlington, MA: Morgan Kaufmann.
- Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis. New York: Springer.
- Yau, N. (2013) Data Points: Visualization That Means Something. Indianapolis, IN: Wiley.

BAB 5

PENERAPAN *DATA ANALYTICS* DI DUNIA NYATA

Oleh Nengah Widya Utami

5.1 Pendahuluan

Data Analytics telah menjadi salah satu pilar utama dalam pengambilan keputusan di berbagai sektor. Dengan kemampuannya mengolah dan menganalisis data dalam jumlah besar, *Data Analytics* membantu organisasi dalam memahami tren, pola, serta memberikan wawasan mendalam yang dapat digunakan untuk meningkatkan efisiensi dan efektivitas kerja. *Data Analytics* melibatkan serangkaian teknik dan alat untuk mengolah data guna mendapatkan wawasan yang bermanfaat. *Data analytics* sendiri merupakan proses mengekstraksi dan menciptakan informasi dari data mentah dengan menyaring, memproses, mengkategorikan, mengkondensasi dan mengkontekstualisasikan data (Wali *et al.*, 2023). Adapun konsep dasar dalam *Data Analytics* meliputi: 1) Pengumpulan Data; 2) Pembersihan Data; 3) Analisis Data; dan 4) Interpretasi Data.

5.2 Penerapan *Data Analytics* pada Berbagai Sektor

Data Analytics memungkinkan analisis data *real-time*, membangun infrastruktur data yang memadai, dan mengoptimalkan berbagai fungsi bisnis seperti penjualan dan pemasaran (Maulani and Widoretno, 2024). Sehingga *Data Analytics* saat ini semakin banyak diterapkan di berbagai industri untuk meningkatkan efisiensi, mengoptimalkan keputusan bisnis, dan menciptakan inovasi baru. Berikut adalah beberapa contoh penerapan *Data Analytics* di berbagai industri:

5.2.1 Bisnis dan Pemasaran

Dalam dunia bisnis, *Data Analytics* digunakan untuk memahami perilaku pelanggan, meningkatkan strategi pemasaran, mengoptimalkan penjualan, dan mendukung inovasi serta pertumbuhan jangka panjang (Prasetya et al., 2024) seperti contoh berikut.

a) Segmentasi pelanggan

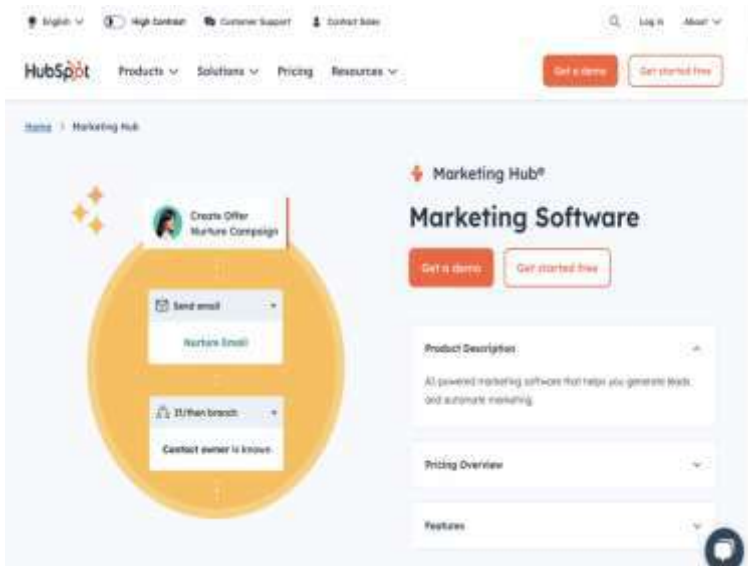
Segmentasi pelanggan memungkinkan perusahaan untuk mengidentifikasi preferensi pelanggan dan menyusun strategi pemasaran yang lebih personal dan efektif (Mukhlis Juliyanto et al., 2024). Beberapa perusahaan yang menggunakan teknologi segmentasi pelanggan untuk mendukung bisnisnya yaitu sebagai berikut.

- E-Commerce (Tokopedia, Shopee, Amazon) menggunakan algoritma clustering seperti K-Means atau DBSCAN untuk mengelompokkan pelanggan berdasarkan pola pembelian, preferensi produk, dan demografi. Misalnya,

pelanggan dapat dikategorikan menjadi “Pembeli Setia,” “Pemburu Diskon,” atau “Pelanggan Baru,” sehingga strategi promosi dapat disesuaikan untuk setiap segmen.

- Bank dan Layanan Keuangan seperti BCA, Mandiri, PayPal menggunakan Customer Lifetime Value (CLV) dan klasifikasi pelanggan berbasis perilaku transaksi untuk menawarkan layanan personal seperti suku bunga khusus atau paket premium kepada pelanggan bernilai tinggi.
- Pariwisata & *Hospitality* (Traveloka, Airbnb) menganalisis preferensi perjalanan pelanggan untuk menawarkan paket wisata yang lebih sesuai, misalnya, wisata budaya untuk segmen yang tertarik dengan sejarah atau wisata petualangan untuk segmen yang lebih muda.

Selain hal tersebut, saat ini terdapat berbagai aplikasi/sistem yang telah dikembangkan dengan fitur segmentasi pelanggan seperti *HubSpot*, *ActiveCampaign*, *SleekFlow*, Google Analytic yang memungkinkan bisnis untuk mengelompokkan pelanggan berdasarkan karakteristik tertentu dan mengirimkan pesan yang dipersonalisasi.



Gambar 5. 1 Aplikasi Segmentasi Pelanggan
(Sumber : <https://www.hubspot.com>)

b) Analisis sentimen

Analisis sentimen merupakan metode untuk mengkaji opini publik sehingga mampu memahami suatu topik dari berbagai perspektif (Hadi Arfian et al., 2025). Adapun industri yang telah memanfaatkan analisis sentimen seperti berikut.

- Industri FMCG (*Coca-Cola, Unilever*) menggunakan *Natural Language Processing* (NLP) untuk menganalisis ulasan pelanggan di platform X, Instagram, dan Facebook guna memahami sentimen terhadap produk baru. Jika banyak komentar negatif, perusahaan dapat segera menyesuaikan strategi pemasaran atau reformulasi produk.

- Industri Hiburan (Netflix, Disney+) menentukan reaksi penonton terhadap film atau serial baru dengan menganalisis ulasan dan komentar di media sosial. Jika banyak sentimen positif terhadap genre tertentu, perusahaan dapat berinvestasi lebih banyak dalam produksi serupa.

Adapun contoh *software* yang dapat digunakan secara *real time* untuk melakukan *Data Analytic* pada ulasan sosial media yaitu *Brand24*, *Hootsuite*, *Brandwatch*, *Google Trend*, dan *MonkeyLearn*.



Gambar 5. 2 Aplikasi Analisis Sentimen

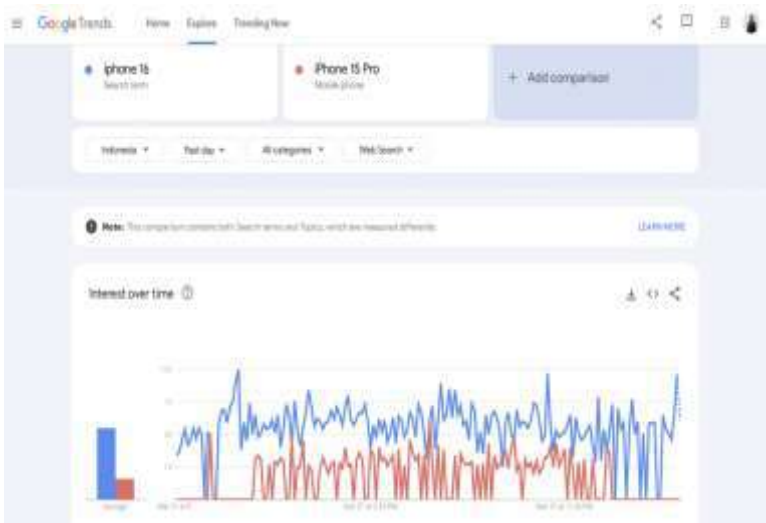
(Sumber : <https://brand24.com>)

c) Prediksi tren pasar

Dengan memanfaatkan volume data yang besar seperti transaksi penjualan, interaksi digital, dan media sosial, perusahaan retail dapat memprediksi tren pasar atau permintaan produk dengan lebih tepat (Halawa, Bangun and Sihombing, 2024). Industri yang telah melakukan prediksi trend yaitu:

- *Retail & Fast-Moving Consumer Goods* (Unilever, P&G) menggunakan model *machine learning* berbasis time-series forecasting (ARIMA, LSTM) untuk memprediksi permintaan produk berdasarkan tren musiman. Contohnya, permintaan produk minuman dingin meningkat di musim panas, sehingga stok harus ditingkatkan sebelum permintaan melonjak.
- *Digital Marketing* (Google Ads, Facebook Ads) memanfaatkan model prediktif untuk menentukan iklan mana yang akan berkinerja terbaik berdasarkan data historis klik, demografi pengguna, dan tren kata kunci. Contohnya, jika tren pencarian produk kecantikan meningkat, maka perusahaan dapat meningkatkan anggaran iklan untuk produk tersebut.
- *Industri Fashion* (Zara, H&M) menggunakan *machine learning* untuk menganalisis data pencarian pelanggan dan media sosial guna memprediksi tren fesyen mendatang. Misalnya, jika warna pastel sedang menjadi tren di Instagram, merek fashion dapat segera menyesuaikan produksi untuk memenuhi permintaan yang akan datang.

Software yang populer yang dapat digunakan untuk prediksi tren pasar seperti *Google Trend*, *Google Cloud AI*, *Facebook Prophet*, dan *Google BigQuery*.



Gambar 5.3 Aplikasi Analisis Tren Pasar
(Sumber : <https://trends.google.com/>)

5.2.2 Kesehatan

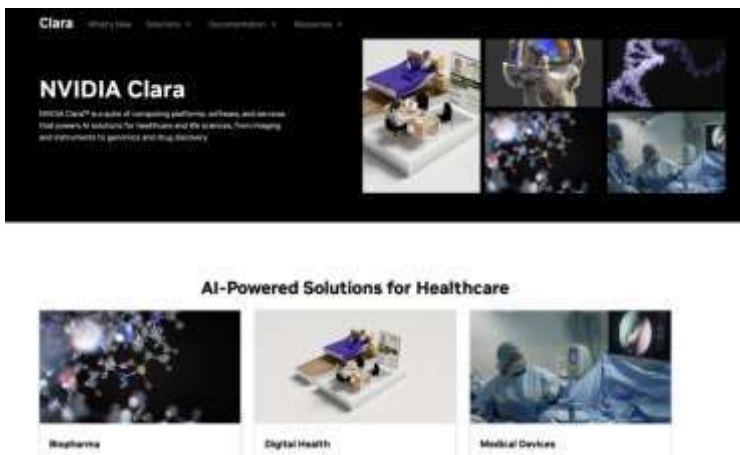
Data perawatan kesehatan yang besar memiliki potensi besar untuk meningkatkan pemahaman tentang efektivitas perawatan di dunia nyata, insiden, manajemen, dan prognosis berbagai kondisi medis terutama untuk penyakit langka (Sulaiman *et al.*, 2023). Dalam industri kesehatan, *Data Analytics* digunakan untuk menganalisis rekam medis pasien, mengoptimalkan perawatan, dan mendukung penelitian medis. Contoh penerapannya meliputi:

a) **Diagnosis penyakit**

Beberapa contoh industri yang memanfaatkan *Data Analytics* untuk diagnosis penyakit sebagai berikut.

- *IBM Watson Health* menggunakan AI untuk menganalisis rekam medis dan membantu diagnosis kanker
- *Google Health* menerapkan AI dalam mendeteksi penyakit mata dan kanker payudara melalui analisis gambar medis
- *Zebra Medical Vision* mengembangkan algoritma AI untuk membaca hasil CT Scan dan MRI guna mendeteksi kelainan lebih cepat

Beberapa tools yang dapat digunakan untuk Data Analytics diagnosis penyakit seperti *NVIDIA Clara*, *Google AutoML Vision*, dan *IBM Watson for Oncology*.



Gambar 5. 4 Aplikasi Diagnosis Penyakit
(Sumber : <https://www.nvidia.com/en-us/clara/>)

b) Prediksi penyebaran penyakit

Berikut adalah beberapa perusahaan yang melakukan prediksi penyebaran penyakit berdasarkan data.

- *BlueDot* (Kanada) pertama kali mendeteksi penyebaran COVID-19 berdasarkan analisis laporan kesehatan global, data penerbangan, dan berita lokal sebelum WHO mengumumkan pandemi.
- *Johns Hopkins University* (JHU) mengembangkan dashboard global untuk melacak penyebaran COVID-19 secara real-time
- *HealthMap* (Harvard Medical School) merupakan sistem pemantauan penyakit yang menggunakan machine learning untuk mengumpulkan dan menganalisis berita, laporan medis, serta media sosial guna mendeteksi wabah penyakit seperti flu, Ebola, dan Zika.



Gambar 5. 5 Aplikasi Prediksi Penyebaran Penyakit
(Sumber : <https://coronavirus.jhu.edu/map.html>)

5.2.3 Keuangan dan Perbankan

Dalam sektor keuangan Data Analytics dapat memberi manfaat bagi auditor seperti mempermudah dalam mengumpulkan bukti, menetapkan populasi yang luas, memprediksi risiko yang ada, dan mempermudah auditor dalam menganalisis data secara cepat. Selain itu, industri keuangan juga memanfaatkan Data Analytics untuk mendeteksi penipuan, mengelola risiko, dan memberikan rekomendasi investasi yang lebih baik. Contohnya:

a) Deteksi transaksi mencurigakan

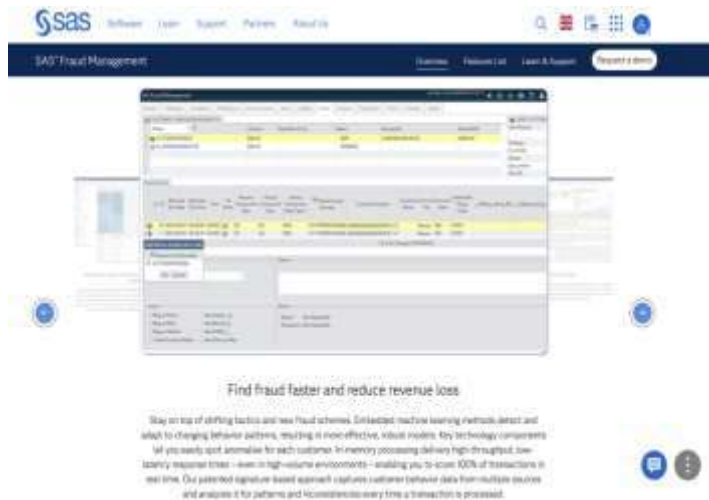
Pada sektor keuangan dan perbankan, beberapa perusahaan telah mengimplementasikan *Data Analytic* untuk mencegah resiko dalam transaksi seperti berikut.

- PayPal menggunakan AI dan machine learning untuk mendeteksi transaksi mencurigakan

dan mencegah fraud dalam pembayaran digital.

- P *Morgan Chase* menggunakan big Data Analytics untuk memantau transaksi real-time dan mendeteksi aktivitas ilegal, seperti money laundering dan identity fraud.
- HSBC menggunakan machine learning untuk mengevaluasi perilaku transaksi nasabah dan mencegah pencucian uang (Anti Money Laundering).

Untuk memproses *big data* transaksi keuangan dalam jumlah besar secara *real-time* dapat menggunakan beberapa platform yang dikembangkan seperti *Hadoop & Spark*, *Splunk*, dan *SAS Fraud Management*.

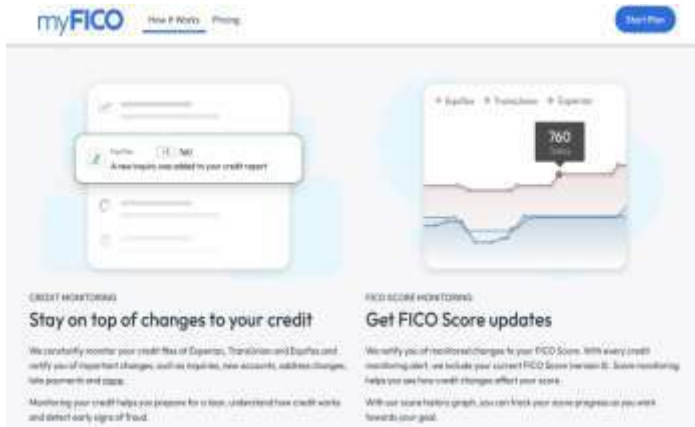


Gambar 5. 6 Aplikasi *Fraud Detection*
(Sumber : <https://www.sas.com/>)

b) Analisis kredit pelanggan

Beberapa perusahaan yang telah memanfaatkan *Data Analytics* dalam analisis kredit, antara lain:

- FICO (*Fair Isaac Corporation*) menggunakan big data analytics untuk menentukan skor kredit pelanggan dan memprediksi risiko gagal bayar.
- *Experian & Equifax* yang merupakan lembaga kredit yang menggunakan machine learning untuk mengevaluasi data finansial pelanggan dan menentukan kelayakan pinjaman.
- *LendingClub & Kabbage* sebuah Fintech yang menggunakan AI untuk menilai kelayakan peminjam berdasarkan data transaksi dan perilaku keuangan.



Gambar 5. 7 Aplikasi Analisis Kredit

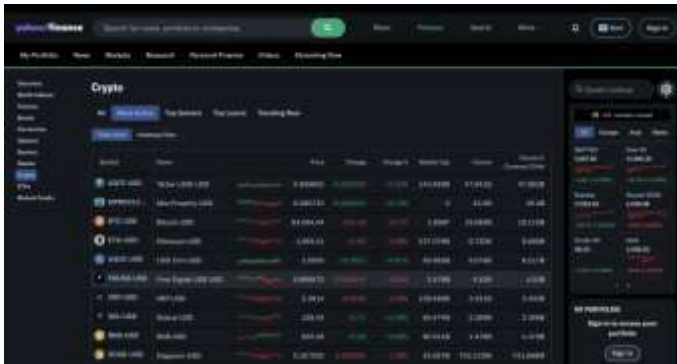
(Sumber : <https://www.myfico.com>)

c) Prediksi tren pasar saham

Beberapa perusahaan yang melakukan prediksi tren saham dengan memanfaatkan Data Analytics yaitu:

- *Bloomberg Terminal* yang menggunakan big data dan AI untuk memprediksi pergerakan saham berdasarkan analisis pasar dan berita global.
- *Goldman Sachs* yang menggunakan *deep learning* dan NLP (*Natural Language Processing*) untuk menganalisis sentimen pasar dari berita keuangan dan laporan perusahaan.
- *Renaissance Technologies* yang merupakan *Hedge fund* yang menggunakan algoritma *machine learning* untuk mengidentifikasi peluang investasi dengan akurasi tinggi.

Tools keuangan yang menyediakan data pasar, berita, dan analisis prediksi saham seperti *Bloomberg Terminal*, *Yahoo Finance API & Alpha Vantage*, dan *AWS SageMaker*.



Gambar 5. 8 Aplikasi Tren Pasar Saham

(Sumber : <https://finance.yahoo.com>)

5.2.4 Pendidikan

Teknologi *Data Analytics* telah membawa perubahan signifikan dalam dunia pendidikan, memungkinkan institusi untuk menyesuaikan metode pengajaran, memberikan rekomendasi kursus yang lebih relevan, serta memprediksi keberhasilan akademik siswa. Berikut adalah beberapa contoh perusahaan yang telah berhasil menerapkan *Data Analytics* dalam pendidikan, beserta platform dan tools yang digunakan.

a) Analisis performa siswa

Contoh implementasi *Data Analytics* untuk analisis performa siswa seperti berikut.

- *Coursera & Udemy* menggunakan machine learning untuk melacak tingkat penyelesaian kursus, mengevaluasi area kelemahan siswa, dan memberikan umpan balik berbasis data kepada instruktur.
- *edX* (dikelola oleh Harvard & MIT) menganalisis keterlibatan siswa di kursus *online*, seperti frekuensi login, durasi menonton video, serta jumlah tugas yang

diselesaikan, untuk memberikan wawasan kepada pengajar tentang efektivitas metode pembelajaran mereka.

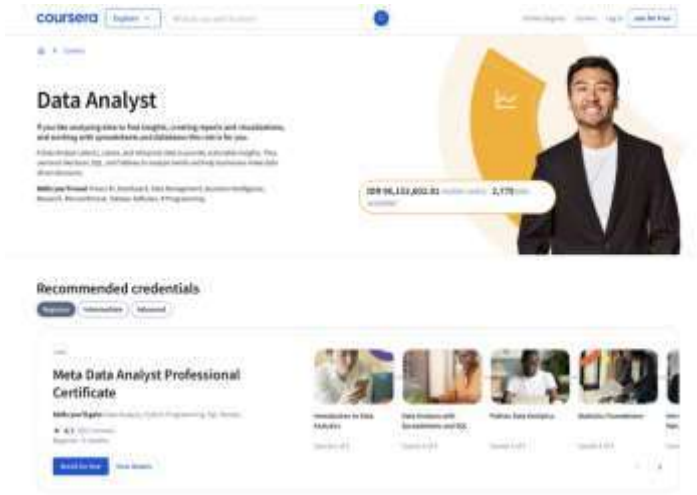


Gambar 5. 9 Aplikasi Analisis Performa Siswa
(Sumber : <https://udemy.com/>)

b) Sistem rekomendasi

- *LinkedIn Learning* yang menggunakan rekomendasi berbasis AI untuk menyarankan kursus yang sesuai dengan riwayat pekerjaan, keterampilan, dan interaksi pengguna di platform.
- *Udacity & Pluralsight* yang menggunakan analitik data untuk memetakan keahlian siswa dan menyesuaikan kursus yang paling sesuai dengan kebutuhan industri. Menganalisis pola belajar pengguna untuk meningkatkan akurasi rekomendasi kursus.
- *Coursera & edX* menggunakan *machine learning* dan NLP untuk menganalisis riwayat pencarian

dan kursus yang pernah diambil siswa, lalu merekomendasikan kursus serupa yang paling cocok.



Gambar 5. 10 Aplikasi Rekomendasi Kursus
(Sumber : <https://www.coursera.org>)

5.2.5 Pemerintahan dan Kebijakan Publik

Pemerintah di berbagai negara telah menerapkan data analytics untuk meningkatkan efisiensi dalam perencanaan transportasi, pengelolaan anggaran, dan keamanan publik. Dengan bantuan *big data*, AI, dan *machine learning*, kebijakan yang diambil menjadi lebih berbasis data dan akurat. Perusahaan yang menerapkan data analytic untuk analisis mobilitas penduduk yaitu:

- *IBM & Smart Cities Projects* yang membantu kota seperti New York, Barcelona, dan Singapura dalam menganalisis pola lalu lintas guna meningkatkan efisiensi transportasi publik dan mengurangi kemacetan.

- *Singapore Land Transport Authority (LTA)* yang menggunakan AI dan machine learning untuk memprediksi kepadatan transportasi publik serta menyesuaikan jadwal MRT dan bus berdasarkan pola mobilitas penduduk.
- *Uber & Lyft* yang berkolaborasi dengan pemerintah untuk menyediakan data *anonymized* tentang perjalanan pengguna guna meningkatkan kebijakan transportasi di perkotaan.



Gambar 5. 11 Aplikasi Perencanaan Transportasi
(Sumber : <https://www.lta.gov.sg>)

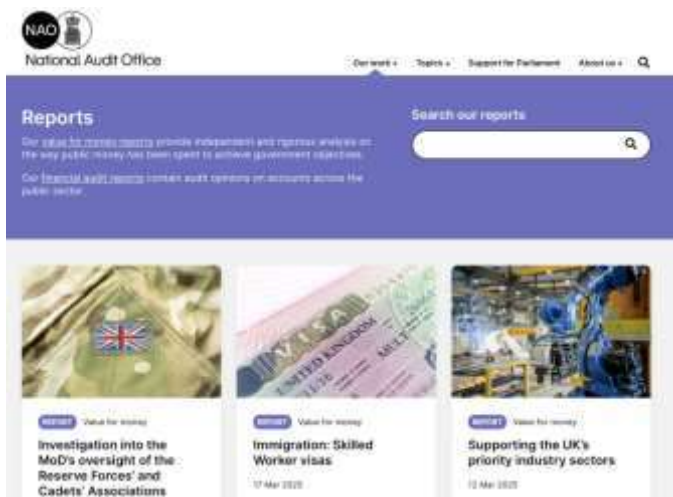
a) Deteksi anomali anggaran

Berikut adalah contoh-contoh industri dan lembaga yang telah menerapkan *Data Analytics* dalam pemerintahan untuk deteksi anomali pengeluaran anggaran negara.

- *U.S. Government Accountability Office (GAO)* yang menggunakan machine learning untuk mendeteksi pengeluaran yang tidak wajar dalam anggaran pemerintah, termasuk fraud

dalam kontrak publik dan proyek infrastruktur.

- UK *National Audit Office* (NAO) yang menggunakan machine learning untuk memverifikasi efisiensi penggunaan anggaran negara dan mengidentifikasi area yang berisiko tinggi terhadap penyalahgunaan dana publik.



Gambar 5. 12 Aplikasi Deteksi Anomali Anggaran Negara

(Sumber : <https://www.nao.org.uk/>)

b) Prediksi tingkat kriminalitas

- *Los Angeles Police Department* (LAPD) menggunakan sistem *PredPol* (*Predictive Policing*) yang berbasis AI untuk memprediksi area yang berisiko tinggi terhadap kejahatan.
- *Interpol & Europol* menggunakan *Data Analytics* untuk menganalisis pola kejahatan

lintas negara dan memprediksi tren kriminalitas internasional.

- *New York Police Department (NYPD)* menggunakan *CompStat (Computer Statistics System)*, yaitu sistem berbasis Data Analytics untuk menganalisis pola kriminalitas dan alokasi sumber daya kepolisian.



Gambar 5. 13 Aplikasi Prediksi Kriminalitas
(Sumber : <https://www.nyc.gov/>)

5.3 Tantangan dalam Implementasi *Data Analytics*

Meskipun memiliki banyak manfaat, penerapan Data Analytics juga menghadapi beberapa tantangan, seperti:

- a) Data sering tidak lengkap, tidak akurat, atau berasal dari berbagai sumber dengan format berbeda, sehingga sulit untuk dianalisis secara optimal.
- b) Risiko kebocoran data dan penyalahgunaan informasi pribadi semakin tinggi, terutama dengan meningkatnya regulasi.
- c) Volume data yang terus bertambah membutuhkan infrastruktur yang mampu menangani penyimpanan dan pemrosesan secara efisien.
- d) Kekurangan tenaga ahli dalam analisis data dan machine learning membuat implementasi Data Analytics menjadi lebih lambat dan kurang optimal.
- e) Model AI dan machine learning dapat memiliki bias yang menyebabkan hasil analisis tidak adil atau kurang akurat dalam pengambilan keputusan.

Untuk mengatasi tantangan ini, perusahaan dan organisasi perlu menerapkan strategi pengelolaan data yang lebih baik, meningkatkan keamanan, serta mengembangkan keahlian dalam analisis data agar penerapan *Data Analytics* dapat berjalan efektif dan memberikan manfaat maksimal.

DAFTAR PUSTAKA

- Hadi Arfian, M. et al. (2025) 'Analisis Sentimen Pada Media Sosial Menggunakan Metode Support Vector Machine', Jurnal Ilmu Teknik dan Komputer, 09(01), pp. 1–6. Available at: <https://doi.org/10.22441/jitkom.v9i1.001>.
- Halawa, P., Bangun, B. and Sihombing, V. (2024) 'Penggunaan Big Data untuk Prediksi Tren Pasar dalam Industri Retail', 4, pp. 32–36. Available at: <https://doi.org/10.55338/justikpen.v4i1.136>.
- Maulani, N.N. and Widoretno, A.A. (2024) 'Analisis Pemanfaatan Data Analytics Dalam Pengambilan Keputusan di PT XYZ', Jurnal Ekonomi Revolusioner, 7(7), pp. 164–170.
- Mukhlis Juliyanto, R. et al. (2024) 'Penerapan Algoritma Clustering Untuk Segmentasi Pelanggan Berdasarkan Perilaku Pembelian', Seminar Nasional Amikom Surakarta (SEMNASA) [Preprint].
- Prasetya, R.P.D. et al. (2024) 'Implementasi Penggunaan Data Analytics untuk Mengoptimalkan Pengambilan Keputusan Bisnis di Era Digital', Jurnal Bisnis dan Komunikasi Digital, 2(2), p. 12. Available at: <https://doi.org/10.47134/jbkdv2i2.3459>.
- Wali, M. et al. (2023) Penerapan & Implementasi Big Data di Berbagai Sektor. PT. Sonpedia Publishing Indonesia. Available at: www.sonpedia.com.
- Sulaiman, T.W. et al. (2023) 'Literature Review: Penerapan Big Data Dalam Kesehatan Masyarakat', 1(3).

Available at: <https://publish.ojs-indonesia.com/index.php/SATUKATA/index>.

BIODATA PENULIS



Ir. Jarot Budiasto, S.T., M.T.

Dosen Program Studi Sistem Informasi
Fakultas Teknik Universitas Musamus

Penulis lahir di Magelang pada 4 Maret 1981. Menyelesaikan pendidikan Sarjana (S1) di Program Studi Teknik Informatika, Universitas Atma Jaya Yogyakarta, pada tahun 2006, dan melanjutkan pendidikan Magister (S2) di bidang Teknik Informatika di universitas yang sama, lulus pada tahun 2008. Pada tahun 2022, penulis memperoleh gelar Insinyur dari Universitas Muslim Indonesia.

Sejak tahun 2008, penulis aktif berkarier sebagai dosen di Program Studi Sistem Informasi, Universitas Musamus, dengan keahlian utama di bidang *Data Analytics* dan *Machine Learning*. Selain berkecimpung di dunia akademik, penulis juga aktif menulis dan telah menerbitkan beberapa buku, di antaranya *Sistem Pendukung Keputusan, Database, Machine Learning untuk Pemula* dan *Python untuk Data Science*.

Melalui karya dan dedikasinya, penulis berkomitmen untuk terus berkontribusi dan berupaya menciptakan dampak positif yang signifikan dalam pengembangan ilmu pengetahuan dan teknologi di Indonesia.

BIODATA PENULIS



Dodi, S.Pd., M.Pd.

Dosen Program Studi Ilmu Kelautan
Fakultas Ilmu Pengetahuan Alam dan Kelautan
Universitas OSO

Penulis berasal dari Padang Tikar dan lahir pada tanggal 20 Agustus 1990. Penulis merupakan seorang dosen tetap di Program Studi Ilmu Kelautan, Fakultas IPA dan Kelautan, Universitas OSO. Ia menyelesaikan pendidikan S1 di program studi Pendidikan Matematika di UNTAN sebelum melanjutkan ke jenjang S2 di Program Studi Pendidikan Matematika di UNTAN. Penulis menekuni dunia kepenulisan pada bidang Matematika dan Sains, serta bidang Pendidikan.

Penulis berharap, karya yang diterbitkannya dapat memberikan inspirasi serta manfaat bagi orang lain untuk belajar matematika dan sains. Kontak penulis tercantum sebagai berikut; email: dodi_a1@yahoo.com / dodi@oso.ac.id; No Hp/ Wa: 089694156039.

BIODATA PENULIS

Ketrin R. Manullang, S.Pd., S.Kom., M.Pd.

Dosen Program Studi Pendidikan Matematika
Fakultas Keguruan dan Ilmu Pendidikan

Penulis lahir di Medan tanggal 13 November 1986. Penulis adalah dosen tetap pada Program Studi Pendidikan Matematika Fakultas Keguruan dan Ilmu Pendidikan. Penulis menyelesaikan pendidikan S1 pada Jurusan Pendidikan Teknik Bangunan dan S1 Teknik Informatika. Penulis kemudian melanjutkan Pendidikan magister S2 pada Jurusan Teknologi Pendidikan. Dalam menjalankan tugasnya sebagai seorang pengajar, penulis juga menekuni bidang menulis dan juga menjadikannya sebagai hobi.

BIODATA PENULIS



Lawrence Adi Supriyono, S.Kom, M.T.

Dosen Program Studi Rekayasa Perangkat Lunak
Fakultas Teknologi Informasi
Universitas Jakarta Internasional

Penulis lahir di Kediri, 31 Mei 1996. Saat ini merupakan Kepala Program Studi & Dosen Tetap di Program Studi Rekayasa Perangkat Lunak, Fakultas Teknologi Informasi, Universitas Jakarta Internasional (UNIJl). Penulis menyelesaikan pendidikan S1 di bidang Sistem Komputer dengan fokus pada dunia Programming dan S2 di bidang Teknik Elektro, dengan fokus pada Mekatronika Digital dan Robotik.

Sebagai akademisi dan peneliti, Penulis memiliki minat utama pada pengembangan perangkat lunak, kecerdasan buatan (AI), machine learning, dan robotika cerdas. Ia telah memimpin berbagai proyek inovatif, termasuk E-SIM (Electronic Simulation System), Sistem Informasi Manajemen Rumah Sakit (SIMRS), serta solusi berbasis teknologi untuk sektor kesehatan, lingkungan, dan edukasi interaktif.

Selain itu, Penulis juga aktif dalam penelitian akademik dengan berbagai publikasi di jurnal nasional maupun internasional, yang semakin mengukuhkan posisinya sebagai peneliti terkemuka di bidangnya. Penulis dapat dihubungi melalui email: lawrence.supriyono@uniji.ac.id.

BIODATA PENULIS



Nengah Widya Utami, S.Pd., M.Kom.

Dosen Program Studi Sistem Informasi Akuntansi
Universitas Primakara

Penulis lahir di Singaraja tanggal 28 Februari 1993. Penulis adalah dosen tetap pada Program Studi Sistem Informasi Akuntansi, Universitas Primakara. Menyelesaikan pendidikan D3 Manajemen Informatika, melanjutkan S1 pada jurusan Pendidikan Teknik Informatika, dan melanjutkan S2 pada jurusan Ilmu Komputer di Universitas Pendidikan Ganesha. Hingga saat ini penulis aktif melakukan pengajaran, penelitian, dan pengabdian kepada masyarakat. Penulis mengampu matakuliah Pemrograman, Sistem Basis Data, dan Data Mining. Sedangkan penelitian berfokus pada Data Science dan Information System. Publikasi yang telah dihasilkan mencakup jurnal maupun prosiding skala nasional dan internasional, HKI, Buku, dan media massa.