

BAB III

PELAKSANAAN KERJA MAGANG

3.1 Kedudukan dan Koordinasi

Posisi *Data Analyst* pada Rumah Sakit Mentari berada di bawah Departemen IT, yang bertanggung jawab atas pengelolaan sistem informasi, infrastruktur teknologi, serta pengolahan data untuk mendukung operasional rumah sakit. Departemen IT memiliki peran yang penting dalam memastikan bahwa semua sistem berjalan dengan efisien, aman, dan data yang dihasilkan dapat digunakan secara maksimal untuk pengambilan keputusan strategis. *Data Analyst* berperan dalam menganalisis data kesehatan dan operasional rumah sakit untuk meningkatkan efisiensi layanan dan mendukung kebijakan manajemen rumah sakit yang berbasis data. Peran ini melibatkan penggunaan teknologi analisis canggih untuk mengolah data yang dapat memberikan *insight* yang diperlukan oleh berbagai pihak dalam organisasi. Dengan cara ini, *Data Analyst* tidak hanya mendukung operasional sehari-hari, tetapi juga berkontribusi dalam perencanaan jangka panjang dan perbaikan berkelanjutan di rumah sakit. Tugas *Data Analyst* mencakup analisis data medis untuk meningkatkan kualitas perawatan pasien, memantau kinerja operasional rumah sakit, serta memberikan rekomendasi berbasis data yang mendukung kebijakan rumah sakit, seperti pengelolaan sumber daya, pengurangan biaya, dan peningkatan pelayanan. Selain itu, *Data Analyst* juga memainkan peran dalam memastikan kepatuhan terhadap regulasi yang ada dengan menyediakan laporan dan analisis yang diperlukan.

Gambar 3.1 menunjukkan struktur hierarki kedudukan *Data Analyst* dalam organisasi Rumah Sakit Mentari. Dalam struktur ini, PT. Mula Sendang Sejahtera berfungsi sebagai badan pengelola utama yang membawahi Direktur Rumah Sakit. Direktur mengawasi berbagai departemen, termasuk Departemen IT, yang menjadi penghubung utama dalam pengelolaan sistem informasi rumah sakit. *Data Analyst*, yang berada di dalam Departemen IT, memiliki koordinasi yang erat dengan tim teknologi informasi lainnya untuk memastikan bahwa data yang dikumpulkan,

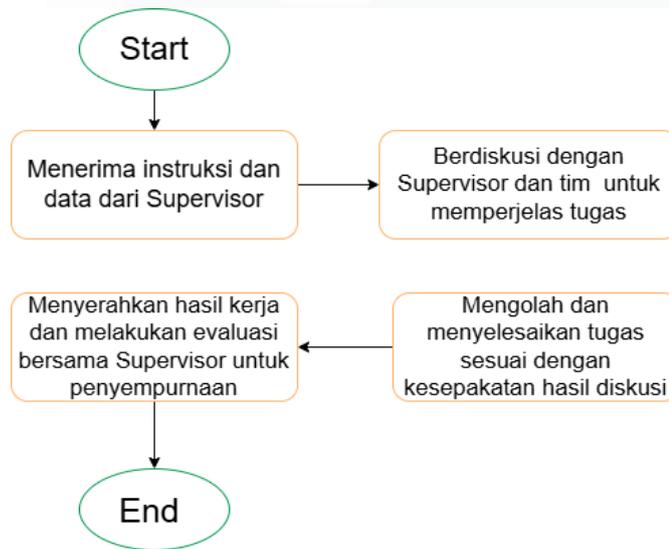
diolah, dan dianalisis dapat memberikan wawasan yang relevan bagi pihak manajemen rumah sakit.



Gambar 3.1 Diagram Hierarki Posisi *Data Analyst*

Sumber: Divisi IT Rumah Sakit Mentari

Dengan struktur ini, *Data Analyst* bekerja secara kolaboratif dengan berbagai divisi untuk memastikan bahwa pengolahan data berjalan dengan baik dan hasil analisisnya dapat dimanfaatkan secara optimal. Kedudukan ini juga memungkinkan adanya komunikasi yang efektif dengan tim teknis, tenaga medis, serta manajemen rumah sakit dalam menerapkan solusi berbasis data. Melalui koordinasi yang jelas, hasil analisis data dapat digunakan untuk meningkatkan layanan pasien, mengoptimalkan efisiensi operasional, serta memberikan rekomendasi strategis bagi rumah sakit dalam pengambilan keputusan berbasis data. Untuk memastikan kelancaran tugasnya, *data analyst* mengikuti alur koordinasi kerja yang terstruktur, sebagaimana ditunjukkan dalam Gambar 3.2. Alur ini menggambarkan tahapan kerja *data analyst* mulai dari menerima instruksi hingga menyerahkan hasil analisis yang telah dievaluasi.



Gambar 3.2 Flow chart Pelaksanaan Magang

Proses kerja *Data Analyst* di Rumah Sakit Mentari dimulai dengan menerima instruksi dan data dari Supervisor, yang mencakup arahan terkait analisis, jenis data yang digunakan, serta tujuan yang ingin dicapai. Setelah menerima instruksi, *Data Analyst* berdiskusi dengan Supervisor dan tim untuk memperjelas tugas, menentukan pendekatan analisis yang sesuai, serta mengidentifikasi kendala yang mungkin dihadapi. Diskusi ini memastikan bahwa data yang digunakan valid dan metodologi analisis yang diterapkan sesuai dengan kebutuhan rumah sakit.

Setelah tahap diskusi, *Data Analyst* mengolah dan menyelesaikan tugas sesuai dengan kesepakatan hasil diskusi. Proses ini meliputi pembersihan data, transformasi, serta penerapan teknik analisis seperti visualisasi data, atau model prediktif. Hasil analisis kemudian disajikan dalam bentuk laporan atau *dashboard* yang dapat digunakan manajemen dalam pengambilan keputusan. Setelah itu, *Data Analyst* menyerahkan hasil kerja dan melakukan evaluasi bersama Supervisor, di mana Supervisor akan meninjau, memberikan umpan balik, serta meminta penyempurnaan jika diperlukan. Dengan mengikuti alur kerja dalam Gambar 3.2, *Data Analyst* dapat memastikan bahwa proses analisis data berjalan secara sistematis dan memberikan manfaat dalam meningkatkan efisiensi operasional serta mendukung strategi manajemen rumah sakit.

3.2 Tugas dan Uraian Kerja Magang

Selama pelaksanaan program magang, peserta magang mengerjakan berbagai tugas yang berhubungan dengan analisis data di Rumah Sakit Mentari. Fokus utama dari kegiatan ini adalah pengolahan dan visualisasi data, yang bertujuan untuk mendukung pengembangan *dashboard* layanan kesehatan di rumah sakit. Untuk mendukung kelancaran pelaksanaan magang sebagai *Data Analyst* di Rumah Sakit Mentari, berbagai tugas telah disusun dalam sebuah jadwal kerja yang sistematis. Berikut merupakan table yang menyajikan rincian kegiatan yang dilakukan selama program magang terdapat pada Tabel 3.1.

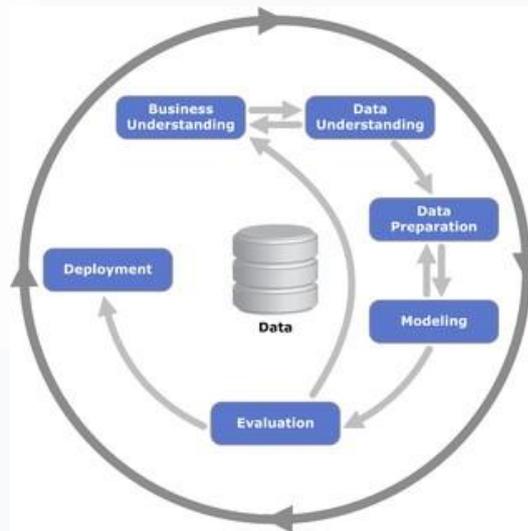
Tabel 3.1 Rincian Kegiatan Pelaksanaan Kerja Magang

No	Kegiatan	Minggu ke-	Tanggal Mulai	Tanggal Selesai
Memahami kebutuhan analisis data (<i>Business Understanding</i>)				
1	Berkolaborasi dengan tim rumah sakit untuk memahami kebutuhan analisis data.	1	7 Januari 2025	10 Januari 2025
2	Identifikasi kebutuhan analisis data dan <i>tools</i> yang akan digunakan.	1	7 Januari 2025	10 Januari 2025
Mengumpulkan, membersihkan, dan memvalidasi data (<i>Data Understanding & Data Preparation</i>)				
3	Mengumpulkan data dari berbagai sumber.	2-9	13 Januari 2024	7 Maret 2025
4	Melakukan <i>data cleaning</i> untuk memastikan kualitas.	2-9	13 Januari 2024	7 Maret 2025
Mengoptimalkan proses Extract, Transform, Load (<i>Data Preparation</i>)				
5	Melakukan eksplorasi dan transformasi data.	2-11	14 Januari 2025	17 Maret 2025

No	Kegiatan	Minggu ke-	Tanggal Mulai	Tanggal Selesai
Menganalisis data menggunakan teknik statistik dan alat analitik (Modeling)				
6	Mengidentifikasi pola tren layanan kesehatan berdasarkan data.	2-11	14 Januari 2025	21 Maret 2025
7	Melakukan prediksi stok obat menggunakan <i>data mining</i> .	12-16	24 Maret 2025	30 April 2025
Mengembangkan visualisasi data dan dashboard interaktif (Modeling)				
8	Mendesain <i>dashboard</i> dan membuat visualisasi menggunakan Power BI.	2-9	13 Januari 2024	7 Maret 2025
Memberikan rekomendasi berbasis data (Evaluation)				
9	Memberikan rekomendasi berdasarkan temuan analisis data dan mengevaluasi hasil analisis dengan tim rumah sakit.	13-16	1 April 2025	30 April 2025
Menyarankan inisiatif perbaikan berdasarkan hasil analisis data (Deployment)				
10	Menyusun laporan akhir serta usulan strategi peningkatan kualitas layanan	16-20	28 April 2025	29 Mei 2025
11	Presentasi akhir hasil pengerjaan tugas	20	30 Mei 2025	30 Mei 2025

Tabel 3.1 merupakan representasi dari rangkaian kegiatan yang telah dilalui oleh peserta magang selama periode magang di Rumah Sakit Mentari. Peserta magang menerapkan metode CRISP-DM sebagai kerangka kerja utama dalam

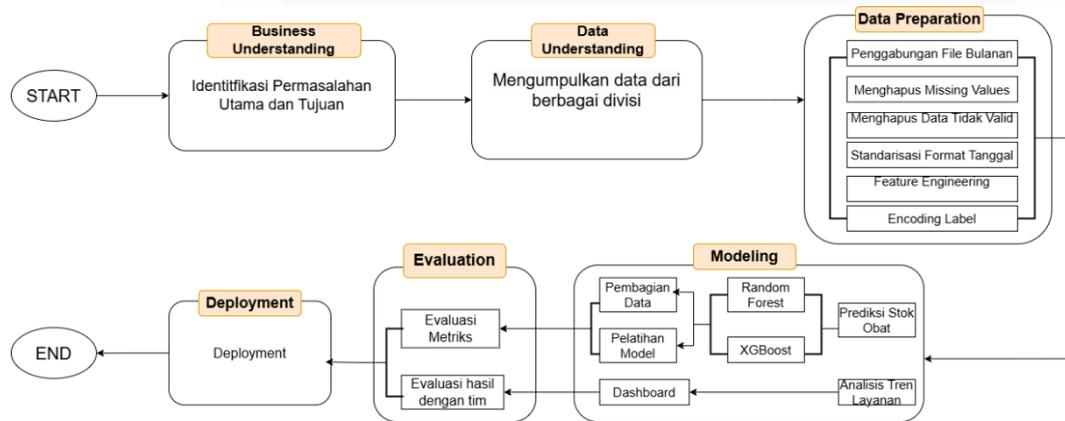
pelaksanaan tugasnya. CRISP-DM (*Cross Industry Standard Process for Data mining*) adalah metodologi standar yang menyediakan alur sistematis dan terstruktur dalam pelaksanaan proyek *data mining*, yang terdiri dari enam tahapan utama: *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation*, dan *Deployment* [12]. Alurnya digambarkan seperti yang ditunjukkan pada Gambar 3.3.



Gambar 3.3 Alur CRISP-DM [13]

Pada tahap pertama *Business Understanding*, peserta magang berkolaborasi dengan tim rumah sakit untuk memahami kebutuhan analisis data yang berkaitan dengan layanan kesehatan, termasuk perencanaan pengembangan *dashboard* interaktif guna mendukung pengambilan keputusan yang cepat dan berbasis data. Pada tahap *Data Understanding*, dilakukan pengumpulan data dari berbagai sumber internal rumah sakit serta eksplorasi awal terhadap kualitas dan struktur data. Tahap *Data Preparation* mencakup proses pembersihan, transformasi, dan pengolahan data untuk memastikan kelayakan data dalam analisis dan visualisasi. Selanjutnya, pada tahap *Modeling*, peserta menggunakan teknik analisis statistik dan *data mining* untuk mengidentifikasi tren layanan dan melakukan prediksi, seperti perencanaan kebutuhan stok obat. Visualisasi data kemudian dikembangkan menggunakan Power BI untuk menghasilkan *dashboard* interaktif yang informatif. Pada tahap *Evaluation*, *dashboard* yang telah dikembangkan dievaluasi melalui diskusi

bersama tim rumah sakit guna memastikan bahwa hasil visualisasi dan analisis mendukung kebutuhan operasional dan strategis. Terakhir, pada tahap *Deployment*, peserta menyusun laporan akhir dan mempresentasikan rekomendasi strategis berbasis data yang dapat diterapkan untuk meningkatkan efisiensi dan kualitas layanan di Rumah Sakit Mentari. Berikut merupakan alur penelitian spesifik yang sesuai dengan framework CRISP-DM yang disajikan pada gambar 3.4.



Gambar 3.4 Alur Penelitian

Diagram ini menekankan hubungan antar fase secara linier dan logis, memperlihatkan bagaimana proses dimulai dari pemahaman masalah hingga implementasi hasil. Setiap blok merepresentasikan aktivitas spesifik yang dilakukan dalam fase tersebut, seperti penggabungan data bulanan, pelatihan model menggunakan *Random Forest* dan *XGBoost*, serta visualisasi hasil dalam bentuk dashboard. Terdapat keterhubungan langsung antara *Modeling* dan *Evaluation* yang menunjukkan adanya proses iteratif menandakan bahwa hasil model tidak langsung digunakan, tetapi terlebih dahulu dievaluasi bersama tim sebelum diterapkan. Selain itu, keberadaan dua output utama dalam *Modeling* (Prediksi Stok Obat dan Analisis Tren Layanan) memperlihatkan fokus ganda dari penelitian ini: efisiensi operasional dan pemahaman tren layanan kesehatan. Secara keseluruhan, diagram ini memberikan gambaran menyeluruh tentang bagaimana setiap tahapan berkontribusi terhadap tujuan akhir penelitian, sekaligus memudahkan pembaca untuk memahami jalannya proses secara sistematis.

Beberapa *tools* yang digunakan dalam melaksanakan tugas selama program magang berlangsung meliputi Power BI, Python, Google Colab, dan Microsoft Excel, berikut penjelasan singkatnya :

1. Power BI

Power BI adalah perangkat lunak analisis data dan visualisasi yang dikembangkan oleh Microsoft. Alat ini memungkinkan pengguna untuk menghubungkan berbagai sumber data, mengolahnya, serta menyajikannya dalam bentuk laporan interaktif dan *dashboard* yang informatif. Dengan antarmuka yang ramah pengguna, Power BI mendukung pemrosesan data dalam skala besar, baik untuk analisis bisnis, pemantauan kinerja, maupun pengambilan keputusan berbasis data [14]. Fitur utamanya mencakup integrasi dengan berbagai database dan layanan cloud, transformasi data menggunakan Power Query, serta pembuatan laporan yang dapat diperbarui secara otomatis. Selain itu, Power BI juga menyediakan fitur DAX (*Data Analysis Expressions*) untuk analisis data yang lebih kompleks serta Power BI Service untuk kolaborasi dan berbagi *dashboard* secara online. Dengan kemampuannya yang komprehensif, Power BI menjadi alat yang sangat penting dalam analisis dan visualisasi data.

2. Python

Python adalah bahasa pemrograman tingkat tinggi yang sering digunakan dalam analisis data, pemodelan statistik, dan pembelajaran mesin [15]. Dengan sintaks yang sederhana dan pustaka yang beragam, Python memungkinkan pengguna untuk melakukan manipulasi data, analisis eksploratif, serta pembuatan model prediktif dengan efisien. Beberapa pustaka utama dalam analisis data menggunakan Python meliputi Pandas untuk manipulasi dan analisis data berbasis tabel, NumPy untuk komputasi numerik yang efisien, Matplotlib dan Seaborn untuk visualisasi data dalam bentuk grafik dan diagram, serta Scikit-learn yang digunakan dalam implementasi algoritma *machine learning* untuk analisis prediktif. Dengan ekosistem yang luas dan fleksibilitas tinggi, Python menjadi salah satu *tools* utama dalam analisis dan pengolahan data.

3. Google Colab

Google Colab adalah layanan berbasis cloud yang memungkinkan pengguna untuk menjalankan kode Python tanpa perlu melakukan instalasi perangkat lunak tambahan [16]. Dengan menggunakan Jupyter Notebook, Google Colab memudahkan pengguna dalam melakukan analisis data, pelatihan model *machine learning*, serta eksplorasi data dalam lingkungan yang kolaboratif. Keunggulan utama dari Google Colab adalah dukungan GPU dan TPU gratis untuk komputasi yang lebih cepat, integrasi dengan Google Drive untuk penyimpanan dan pengelolaan dataset, kemudahan dalam berbagi kode dan hasil analisis dengan tim, serta dukungan terhadap pustaka Python yang populer dalam analisis data dan *machine learning*. Karena fleksibilitas dan skalabilitasnya, Google Colab menjadi alat yang banyak digunakan oleh data scientist dan analis data dalam eksplorasi serta pemrosesan data berbasis cloud.

4. Microsoft Excel

Microsoft Excel adalah aplikasi perangkat lunak yang digunakan untuk mengelola, menganalisis, dan menyajikan data dalam format tabel. Sebagai bagian dari paket Microsoft Office, Excel memiliki berbagai fitur yang mendukung analisis data, mulai dari perhitungan matematis hingga pembuatan grafik dan *dashboard* sederhana [17]. Fungsi utama Excel mencakup penggunaan formula dan fungsi seperti *SUM*, *AVERAGE*, *VLOOKUP*, dan *IF* untuk perhitungan otomatis, serta PivotTable yang memudahkan analisis dan penyaringan data yang lebih kompleks. Selain itu, fitur *Conditional Formatting* memungkinkan pengguna menyoroti pola atau tren dalam data, sementara *Data Visualization* membantu dalam pembuatan grafik dan diagram untuk menyajikan data secara visual. Excel tetap menjadi alat yang penting dalam pengelolaan data karena kemudahan penggunaannya dan fleksibilitasnya dalam berbagai analisis bisnis dan keuangan.

Alat-alat di atas mendukung keberlangsungan program magang yang dimulai secara resmi sejak tanggal 7 Januari 2025 dan berlangsung selama kurang lebih 5

bulan. Seiring dengan penyusunan laporan ini, program magang masih dalam proses keberlangsungan hingga tanggal yang telah ditentukan, yaitu tepat pada tanggal 31 Mei 2025. Dalam pelaksanaannya, alat-alat ini telah dipercaya untuk membantu meningkatkan kualitas data serta optimalisasi sistem informasi rumah sakit.

3.2.1 Memahami kebutuhan analisis data

Sebelum memulai proses analisis data, langkah paling fundamental adalah memahami terlebih dahulu kebutuhan bisnis atau *business understanding*. Tahapan ini sangat krusial karena berperan sebagai pondasi dalam keseluruhan siklus analitik. Kesalahan dalam memahami kebutuhan pada tahap awal akan berdampak pada tidak relevannya hasil analisis yang dihasilkan, sehingga proses selanjutnya menjadi kurang efektif bahkan bisa menjadi sia-sia. Oleh karena itu, tahap memahami kebutuhan analisis data bertujuan untuk menggali konteks operasional, tantangan utama yang dihadapi, serta sasaran strategis yang ingin dicapai oleh organisasi, dalam hal ini adalah rumah sakit.

Pemahaman terhadap kebutuhan analisis data dilakukan melalui pendekatan kolaboratif antara tim analis data dengan perwakilan dari berbagai divisi rumah sakit. Kolaborasi ini melibatkan diskusi, wawancara, hingga observasi langsung terhadap proses kerja dan sistem informasi yang digunakan oleh masing-masing divisi. Tujuan utama dari pendekatan ini adalah untuk mengidentifikasi informasi apa saja yang dibutuhkan oleh setiap divisi, bagaimana struktur data yang tersedia saat ini, serta bagaimana data tersebut dapat membantu menjawab pertanyaan-pertanyaan bisnis yang relevan dan strategis.

1. Berkolaborasi dengan Tim Rumah Sakit untuk Memahami Kebutuhan Analisis Data

Pada tahap ini, setiap divisi rumah sakit dijadikan mitra aktif dalam proses identifikasi kebutuhan analitik. Masing-masing divisi tentu memiliki fokus, tantangan, dan sasaran yang berbeda. Misalnya, Divisi

Teknologi Informasi (IT) memiliki kebutuhan untuk memantau efektivitas layanan IT internal rumah sakit. Hal-hal yang menjadi perhatian mereka meliputi unit mana saja yang paling sering menggunakan layanan IT, jenis permasalahan teknis yang sering dilaporkan, serta waktu respons dari tim IT dalam menangani kendala tersebut. Melalui analisis ini, tim IT dapat mengevaluasi kinerja layanan mereka dan merencanakan perbaikan berbasis data.

Divisi Farmasi, di sisi lain, memiliki kebutuhan untuk memahami performa penjualan obat. Selain itu, mereka juga membutuhkan alat bantu untuk memprediksi kebutuhan obat di masa depan guna menghindari terjadinya *overstock* atau *stockout*. Divisi Laboratorium fokus pada distribusi jenis pemeriksaan medis, tren permintaan layanan berdasarkan jenis tes, dan pola kunjungan pasien rawat jalan serta rawat inap. Sementara itu, Divisi *General Affair* membutuhkan informasi operasional seperti efektivitas distribusi pekerjaan staf, ketepatan waktu pengambilan limbah medis, distribusi linen rumah sakit, serta kepatuhan staf terhadap penggunaan Alat Pelindung Diri (APD).

Pendekatan kolaboratif ini tidak hanya bermanfaat dalam menggali kebutuhan, tetapi juga membantu membangun pemahaman bersama tentang bagaimana data dapat dimanfaatkan secara optimal. Setiap masukan dari divisi terkait menjadi dasar untuk menentukan indikator kinerja utama, struktur data yang dibutuhkan, serta teknik analisis yang relevan.

2. Identifikasi Kebutuhan Analisis Data dan *Tools* yang Akan Digunakan

Setelah memperoleh pemahaman yang komprehensif mengenai kebutuhan informasi dari setiap divisi, langkah selanjutnya adalah menentukan pendekatan teknis dan *tools* yang akan digunakan untuk mewujudkan proses analisis data yang efektif dan efisien. Dua *tools* utama yang dipilih dalam proyek ini adalah Python dan Power BI. Python dipilih karena fleksibilitasnya dalam menangani proses *data cleaning*, analisis

statistik, serta penerapan model prediksi yang kompleks. Selain itu, Python memungkinkan integrasi otomatis dengan berbagai format data yang digunakan oleh rumah sakit, seperti file Excel atau CSV.

Sementara itu, Power BI digunakan untuk menyusun *dashboard* visual interaktif yang menampilkan hasil analisis dalam bentuk grafik, tabel, dan indikator performa secara dinamis. Keunggulan Power BI terletak pada kemampuannya untuk menyajikan data secara real-time, memperbarui visualisasi secara otomatis, serta memberikan kemudahan dalam eksplorasi data oleh pengguna non-teknis. Dengan demikian, hasil analisis yang telah diolah dengan Python dapat disajikan dengan cara yang lebih komunikatif dan mudah dipahami oleh manajemen dan staf rumah sakit. Penggunaan dua *tools* ini secara terpadu memberikan keuntungan dari sisi kecepatan pengolahan data, fleksibilitas dalam eksplorasi analisis, serta kemudahan dalam penyampaian informasi. Dengan adanya pemilihan *tools* yang tepat sejak awal, proses analisis data dapat dijalankan secara konsisten dan mendalam, serta dapat disesuaikan dengan kebutuhan masing-masing divisi secara berkelanjutan.

3.2.2 Mengumpulkan, membersihkan, dan memvalidasi data

Tahap ini mencakup proses pengumpulan data dari berbagai sumber, pembersihan data, serta validasi untuk menjamin kualitas data yang akan digunakan dalam analisis. Dalam konteks kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data mining*), proses ini berada pada dua fase utama, yaitu *Data Understanding* dan *Data Preparation*. Pada fase *Data Understanding*, fokus utamanya adalah memperoleh pemahaman awal terhadap data yang tersedia, mengidentifikasi struktur data, kualitas data, serta potensi masalah yang mungkin timbul. Data dikumpulkan dari berbagai divisi rumah sakit, seperti Divisi IT yang menyediakan data operasional sistem informasi, Divisi Farmasi yang menyajikan data stok obat, distribusi resep, dan pengadaan farmasi, Divisi Laboratorium dengan data hasil pemeriksaan medis dan layanan diagnostik, serta Divisi *General Affair* yang memiliki data terkait logistik, fasilitas, dan kepegawaian.

Setelah proses pemahaman dan pengumpulan data dilakukan, tahap berikutnya adalah *Data Preparation*, yaitu fase di mana data dibersihkan agar siap dianalisis. Pembersihan ini penting agar data menjadi seragam, dapat diolah secara sistematis, dan tidak menghasilkan bias dalam proses analisis. Validasi dilakukan sebagai langkah akhir untuk memastikan bahwa data yang telah dibersihkan sesuai dengan standar kualitas.

1. Mengumpulkan Data dari Berbagai Sumber

Proses pengumpulan data dilakukan dengan bekerja sama dengan masing-masing divisi untuk mendapatkan data yang dibutuhkan sesuai dengan tujuan analisis. Divisi IT menyediakan data terkait laporan keluhan teknis, waktu respons, durasi penyelesaian masalah, serta jenis perangkat yang paling sering bermasalah. Divisi Farmasi memberikan data mengenai jumlah dan jenis obat yang terjual, stok yang tersedia, serta tren pembelian oleh pasien dan unit terkait. Divisi Laboratorium menyajikan data hasil pemeriksaan medis, termasuk distribusi jenis pemeriksaan, kategori layanan, serta tren permintaan tes laboratorium. Sementara itu, Divisi *General Affair* menyuplai data operasional seperti distribusi pekerjaan, kepatuhan penggunaan APD, ketepatan pengambilan limbah, dan tingkat respons terhadap permintaan layanan.

Setelah data dikumpulkan dari berbagai sumber, data tersebut biasanya tersimpan dalam format Excel, dengan beberapa sheet yang berisi informasi spesifik. Data dari masing-masing divisi umumnya tersimpan dalam format Excel, dengan setiap sheet mewakili jenis informasi tertentu. Salah satu contoh dataset yang digunakan dalam analisis adalah data penjualan obat dari Divisi Farmasi. Data ini mencakup berbagai atribut seperti kode obat, nama obat, jumlah unit terjual, harga satuan, total penjualan, dan tanggal transaksi. Informasi tersebut digunakan untuk mengidentifikasi tren penjualan berdasarkan periode waktu tertentu, mengetahui obat dengan permintaan tertinggi, serta membantu dalam pengelolaan stok agar tetap seimbang dan efisien. Pada tahap awal, data yang diterima dari Divisi Farmasi berjumlah 96.344 baris, Namun, setelah

dilakukan proses *data cleaning* maka jumlah data yang dapat dianalisis menjadi 972 baris. Data yang tersisa inilah yang dianggap valid dan representatif untuk dianalisis lebih lanjut.

Dengan memahami data dari berbagai divisi, termasuk data penjualan obat, perusahaan dapat mengoptimalkan strategi operasional dan pengambilan keputusan berbasis data. Analisis pola penjualan memungkinkan manajemen mengantisipasi lonjakan permintaan, mengelola stok dengan lebih efisien, serta mengevaluasi efektivitas strategi pemasaran dan harga. Selain itu, integrasi data antar divisi juga memungkinkan identifikasi hubungan antarproses, seperti keterkaitan antara tren penjualan dan hasil layanan medis atau keluhan pasien. Pendekatan ini mendorong terciptanya sistem informasi yang saling terhubung, yang dapat mempercepat proses evaluasi dan respons terhadap permasalahan layanan. Dengan demikian, rumah sakit dapat menerapkan strategi berbasis bukti yang lebih tepat sasaran dan berdampak langsung pada kualitas layanan kepada pasien.

Sebagai dasar dari analisis lanjutan, penting untuk memahami bentuk dan struktur data asli yang diterima dari divisi terkait. Oleh karena itu, untuk memberikan gambaran lebih jelas mengenai format awal dari dataset yang digunakan, Tabel 3.2 menyajikan data mentah penjualan obat secara rinci. Tabel ini menampilkan berbagai atribut penting yang terekam dalam setiap transaksi penjualan, seperti tanggal transaksi, nama obat, satuan, jumlah yang terjual, nilai subtotal, biaya, profit, nama pasien, hingga nama dokter yang menangani. Penyajian data mentah dalam Tabel 3.2 bertujuan untuk memperlihatkan kompleksitas dan kekayaan informasi yang tersedia sebelum dilakukan proses pembersihan dan transformasi data. Dengan memahami isi tabel ini, pembaca dapat memperoleh gambaran awal mengenai kondisi dan kualitas data yang digunakan sebagai dasar dalam proses analisis yang lebih lanjut.

Tabel 3.2 Data Mentah Penjualan Obat

Tanggal	Departemen	Gudang	User	Kode Item	Nama Item	Satuan	Jumlah	Subtotal	Cost	Profit	No. Reg	Penjamin	Nama Pasien	Dokter	No Resep	Pabrik
2024-01-01	ICU	INSTALASI FARMASI	Sita Dwi A	OBT000179	NORAXIN	SYRINGE	1	319347	168000	153864	2.31E+09	UMUM	SIR ARY Setiawan M	Michael Y, dr. SpPD	R240 108000001	DIPA
2024-01-01	ICU	INSTALASI FARMASI	Sita Dwi A	OBT000526	NACL 0.9% 100 ML	PLABOT	4	83944	25676	58268	2.31E+09	UMUM	SIR ARY Setiawan M	Michael Y, dr. SpPD	R240 108000001	OTSUKA
2024-01-01	ICU	INSTALASI FARMASI	Sita Dwi A	OBT000013	NACL 0.9% 25 ML	NEBULE	1	9614	3334	6280	2.31E+09	UMUM	SIR ARY Setiawan M	Michael Y, dr. SpPD	R240 108000001	OTSUKA
2024-01-01	ICU	INSTALASI FARMASI	Sita Dwi A	OBT00112	BISOLVON 10 MG/5 ML (50 ML) SOLUTION	BOTOL	1	145987	79773	66214	2.31E+09	UMUM	SIR ARY Setiawan M	Michael Y, dr. SpPD	R240 108000002	SANOFI
2024-01-01	ICU	DEPO FARMASI	Meitria Vesyaka Dewi	OBT01432	VIVENA-9 INJ	VIAL	1	101648	56088	45560	2.31E+09	UMUM	SIR ARY Setiawan M	Kevin Dermawan, dr	I240 108400001	DEXA MEDICA
2024-01-01	ICU	INSTALASI FARMASI	DEWI WULANDARI	OBT01269	WATER FOR INJEKTION 25ML OTSU	PLABOT	4	38456	10654	27802	2.31E+09	UMUM	SIR ARY Setiawan M	Michael Y, dr. SpPD	R240 108000003	OTSUKA

2. Melakukan *Data cleaning* untuk Memastikan Kualitas

Setelah data terkumpul, tahap penting yang perlu dilakukan adalah proses *data cleaning* untuk memastikan bahwa data yang digunakan dalam analisis bersih dan layak. Proses ini bertujuan untuk menghilangkan entri yang tidak valid, tidak lengkap, atau tidak relevan, yang dapat memengaruhi akurasi hasil analisis. Dalam konteks ini, data yang dibersihkan berasal dari laporan bulanan penjualan obat dan alat kesehatan. Dengan menggunakan bahasa pemrograman Python, proses dilakukan secara otomatis untuk seluruh bulan dari Januari hingga Desember, dengan membaca setiap file Excel, menyaring data berdasarkan kondisi tertentu, lalu menyimpannya kembali dalam format baru.

```
In [1]: import os
import pandas as pd

In [2]: this_folder = os.getcwd()

for month in range(1, 13):
    input_file = f"Laporan_Penjualan_Obat_dan_Alkes_{month:02d}-24.xlsx"
    input_path = os.path.join(this_folder, input_file)

    if os.path.exists(input_path):
        print(f"Processing: {input_file}")
        df = pd.read_excel(input_path)

        def remove_racikan(df):
            df = df[~((df['Fabrik'] == "LAINLAIN") | (df['Tanggal'].isna()) | ((df['Subtotal'] == 0)))]
            return df

        df = remove_racikan(df)

        output_file = f"Laporan_Penjualan_Obat_dan_Alkes_{month:02d}-24.xlsx"
        output_path = os.path.join(this_folder, output_file)

        df.to_excel(output_path, index=False)
        print(f"Saved: {output_file}")
    else:
        print(f"File not found: {input_file}")

Processing: Laporan_Penjualan_Obat_dan_Alkes 01-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 01-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 02-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 02-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 03-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 03-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 04-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 04-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 05-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 05-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 06-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 06-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 07-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 07-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 08-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 08-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 09-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 09-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 10-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 10-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 11-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 11-24.xlsx
Processing: Laporan_Penjualan_Obat_dan_Alkes 12-24.xlsx
Saved: Laporan_Penjualan_Obat_dan_Alkes 12-24.xlsx
```

Gambar 3.5 Proses *Data cleaning*

Sebagaimana ditunjukkan pada Gambar 3.6, kode Python tersebut menghapus baris-baris data yang mengandung nilai yang tidak sesuai, seperti jika kolom 'Pabrik' bernilai "LAINLAIN", kolom 'Tanggal' tidak diisi (missing), atau nilai 'Subtotal' sama dengan nol. Baris-baris ini dianggap tidak layak dianalisis karena tidak menyajikan informasi bermakna atau dapat menyesatkan kesimpulan. Proses ini memastikan

bahwa hanya data yang valid dan bersih yang digunakan dalam analisis selanjutnya. Hasilnya adalah file yang telah diperbarui dan disimpan kembali untuk setiap bulan, yang menandakan bahwa proses pembersihan data telah dilakukan secara sistematis dan efisien untuk mendukung kualitas pengambilan keputusan berbasis data di rumah sakit.

Langkah selanjutnya setelah *data cleaning* adalah melakukan validasi hasil pembersihan data. Validasi ini penting untuk memastikan bahwa proses filtering tidak menghapus data yang seharusnya dipertahankan. Misalnya, perlu dicek apakah ada kemungkinan nilai “LAINLAIN” pada kolom 'Pabrik' sebenarnya valid dalam konteks tertentu atau apakah ada kesalahan input data yang seharusnya dikoreksi, bukan langsung dihapus. Oleh karena itu, tinjauan manual atas sebagian data hasil cleaning bisa menjadi langkah tambahan sebelum proses analisis dimulai.

Selain validasi, proses pembersihan ini juga memungkinkan standarisasi data. Misalnya, dalam kasus lain yang serupa, proses *data cleaning* bisa diperluas untuk menyamakan format tanggal, menyatukan ejaan yang berbeda untuk nama pabrik yang sama, serta memastikan bahwa nilai numerik seperti 'Subtotal' berada dalam rentang yang wajar. Standarisasi ini krusial agar pengolahan dan analisis data tidak menghadapi hambatan akibat inkonsistensi data.

Implementasi skrip Python secara otomatis untuk semua bulan menunjukkan efisiensi dan reusabilitas kode dalam pengolahan data rutin. Dengan pendekatan seperti ini, rumah sakit atau organisasi kesehatan dapat menghemat waktu signifikan dan meminimalkan kesalahan manusia yang sering terjadi pada proses manual. Otomatisasi juga memungkinkan integrasi lebih mudah dengan sistem pelaporan atau *dashboard* analitik yang dapat diperbarui secara berkala.

Lebih jauh, proses *data cleaning* ini merupakan bagian integral dari pipeline analisis data yang lebih luas, mulai dari ekstraksi, transformasi, hingga visualisasi dan pelaporan. Dengan memastikan bahwa data sudah bersih pada tahap awal, analisis selanjutnya seperti tren penjualan bulanan,

identifikasi produk paling laku, hingga efisiensi distribusi dapat dilakukan dengan lebih akurat dan terpercaya.

Pembersihan data yang sistematis dan terdokumentasi dengan baik memberikan nilai tambah dalam hal auditabilitas dan replikasi. Artinya, jika di kemudian hari terjadi perbedaan hasil analisis atau perlu dilakukan penelusuran ulang terhadap sumber data, proses pembersihan yang terdokumentasi ini dapat menjadi acuan. Hal ini sangat penting dalam lingkungan seperti rumah sakit yang membutuhkan akuntabilitas tinggi dalam pengelolaan data dan pengambilan keputusan berbasis informasi.

3.2.3 Mengoptimalkan Proses ETL (*Extract, Transform, Load*)

Bagian ini merupakan bagian dari tahap *Data Understanding* dalam metodologi CRISP-DM (*Cross Industry Standard Process for Data mining*). Pada tahap ini, peneliti atau analis bertujuan untuk mengenal struktur, kualitas, dan konteks data yang tersedia agar dapat memahami pola yang tersembunyi dan menyiapkannya secara optimal untuk analisis lebih lanjut. Dalam konteks rumah sakit, data yang digunakan sering kali berasal dari berbagai sistem informasi, dan masing-masing memiliki format serta kualitas data yang berbeda-beda. Oleh karena itu, proses ETL (*Extract, Transform, Load*) menjadi sangat krusial karena berfungsi untuk mengekstrak data dari sumber-sumber tersebut, membersihkannya dari inkonsistensi atau kesalahan, mentransformasikannya menjadi format yang dapat dianalisis, dan memuatnya ke dalam sistem atau file yang siap digunakan dalam tahap *Modeling* dan visualisasi.

Pada tahap ini dilakukan proses *Extract, Transform, Load (ETL)* yang sangat penting untuk memastikan bahwa data siap digunakan untuk analisis. Seperti pada Gambar 3.7, proses dimulai dari *Extract*, yaitu mengambil data dari file Excel bernama *LaporanPenjualanObat.xlsx*. File ini berisi data penjualan obat termasuk tanggal transaksi, kode item, dan jumlah terjual. Data tersebut kemudian diimpor ke dalam Python menggunakan pustaka *pandas*. Setelah data berhasil diekstrak, tahap

Transform dilakukan untuk membersihkan dan merapikan struktur data. Pertama, kolom "Tanggal" dikonversi menjadi format datetime agar dapat diolah berdasarkan waktu. Kemudian dibuat kolom baru bernama "Bulan" dengan format "Nama Bulan Tahun" seperti "Januari 2024" agar memungkinkan data dikelompokkan dengan lebih mudah berdasarkan periode bulanan.

```
In [17]: # import pandas as pd
df = pd.read_excel('LaporanPenjualanObat.xlsx')

In [18]: # memastikan kolom tanggal berupa tipe datetime
df['Tanggal'] = pd.to_datetime(df['Tanggal'])

In [19]: # buat kolom "Bulan" dalam format "Month Year"
df['Bulan'] = df['Tanggal'].dt.strftime('%B %Y')

In [20]: # pivot data agar Kode Item jadi baris dan Bulan jadi kolom
pivot_df = df.pivot_table(
    index='Kode Item',
    columns='Bulan',
    values='Jumlah',
    aggfunc='sum',
    fill_value=0
)

In [10]: # reset index
pivot_df = pivot_df.reset_index()

Bulan Kode Item April 2024 August 2024 December 2024 February 2024 \
0 OBT00001 0 0 0 0
1 OBT00004 15 21 295 131
2 OBT00006 95 111 50 0
3 OBT00008 14 40 0 5
4 OBT00010 4 2 0 0

Bulan January 2024 July 2024 June 2024 March 2024 May 2024 \
0 0 0 0 0 0
1 35 24 61 10 124
2 120 63 0 38 40
3 30 53 45 -2 37
4 1 0 0 0 2
```

Gambar 3.6 Proses ETL

```
In [12]: # urutkan kolom bulan sesuai urutan kalender
bulan_cols = pivot_df.columns[1:]

In [13]: # ubah ke tipe datetime untuk diurutkan
bulan_sorted = sorted(bulan_cols, key=lambda x: pd.to_datetime(x, format='%B %Y'))

In [15]: # susun ulang kolom: KodeBarang + bulan yg sudah diurutkan
pivot_df = pivot_df[['Kode Item'] + bulan_sorted]
print(pivot_df.head())

Bulan Kode Item January 2024 February 2024 March 2024 April 2024 \
0 OBT00001 0 0 0 0
1 OBT00004 35 131 10 15
2 OBT00006 120 0 38 95
3 OBT00008 30 5 -2 14
4 OBT00010 1 0 0 4

Bulan May 2024 June 2024 July 2024 August 2024 September 2024 \
0 0 0 0 0 0
1 124 61 24 21 181
2 40 0 63 111 90
3 37 45 53 40 15
4 2 0 0 2 0

Bulan October 2024 November 2024 December 2024
0 0 14 0
1 171 413 295
2 180 160 50
3 0 30 0
4 0 0 0

In [16]: # Simpan hasil ke file Excel baru
pivot_df.to_excel('LaporanPenjualan_Pivot.xlsx', index=False)
```

Gambar 3.7 Proses Lanjutan ETL

Langkah berikutnya adalah membuat pivot table yang mengubah struktur data agar lebih ringkas dan terstruktur. Dalam pivot ini, "Kode Item" dijadikan baris, sementara "Bulan" dijadikan kolom. Nilai yang

ditampilkan adalah total penjualan untuk setiap item di tiap bulan. Proses ini mempermudah pembacaan tren penjualan secara periodik. Agar kolom bulan tersusun sesuai urutan kalender, nama-nama bulan diubah kembali ke format *datetime* lalu diurutkan. Hal ini mencegah kesalahan seperti bulan “Oktober” muncul sebelum “Februari”. Setelah itu, data disusun ulang dengan urutan kolom yang benar dan rapi. Hasil akhirnya menunjukkan ringkasan penjualan setiap item obat per bulan. Langkah terakhir dalam proses ETL adalah *Load*, yaitu menyimpan hasil transformasi ke dalam file Excel baru. File tersebut dinamai *LaporanPenjualan_Pivot.xlsx* untuk membedakan dengan file awal. File ini siap digunakan untuk analisis lanjutan, visualisasi data, atau pelaporan. Berikut merupakan penjelasan urutan proses ETL secara rinci di dalam Tabel 3.3.

Tabel 3.3 Rincian Proses ETL

Tahapan ETL	Langkah Teknis	Hasil
<i>Extract</i>	Membaca data dari <i>LaporanPenjualanObat.xlsx</i> menggunakan <i>pandas</i>	Data mentah berisi: kolom Tanggal, Departemen, Gudang, User, Kode Item, Nama Item, Satuan, Jumlah, Cost, Profit, Penjamin, Nama Pasien, Nama Dokter, Nomor Resep, Pabrik.
<i>Transform 1</i>	Mengubah kolom Tanggal ke format <i>datetime</i>	Kolom tanggal kini dikenali sebagai format waktu yang bisa diurutkan
<i>Transform 2</i>	Menambahkan kolom baru Bulan (format: Nama Bulan Tahun)	Data memiliki atribut waktu dalam bentuk agregat bulanan.
<i>Transform 3</i>	Membuat pivot tabel: Kode Item sebagai baris, Bulan sebagai kolom, dan nilai adalah jumlah total terjual	Data terstruktur per item dan per bulan memudahkan identifikasi tren.
<i>Transform 4</i>	Mengurutkan kolom bulan berdasarkan urutan kalender	Kolom bulan tersusun Januari-Desember.
<i>Load</i>	Menyimpan hasil transformasi ke file baru <i>LaporanPenjualan_Pivot.xlsx</i>	File siap pakai untuk visualisasi dan analisis lanjutan.

Tampilan perubahan data yang dihasilkan dari proses ETL adalah sebagai berikut di Tabel 3.4. Data yang awalnya bersifat mentah dan tersebar kini telah terorganisir dalam format pivot yang ringkas dan

informatif. Struktur tabel menjadi lebih intuitif, dengan setiap baris mewakili satu item obat dan setiap kolom menunjukkan total penjualan per bulan dalam satu tahun. Format ini memudahkan proses analisis tren penjualan secara periodik serta mempercepat pengambilan keputusan berbasis data. Selain itu, penataan kolom berdasarkan urutan bulan yang sesuai kalender membantu menghindari kekeliruan interpretasi waktu, sehingga hasil analisis menjadi lebih akurat dan relevan.

Tabel 3.4 Beberapa Sample Data Hasil Pre-Processing Menggunakan Metode ETL

Kode Item	Jan 2024	Feb 2024	Mar 2024	April 2024	May 2024	June 2024	Juli 2024	Agustus 2024	Sept 2024	Okt 2024	Nov 2024	Des 2024
OBT00001	0	0	0	0	0	0	0	0	0	0	14	0
OBT00004	35	131	10	15	124	61	24	21	181	171	413	295
OBT00006	120	0	38	95	70	0	63	111	90	180	160	50
OBT00008	30	5	3	14	3	45	53	40	15	0	30	0

Transformasi dalam ETL mencakup beberapa teknik seperti normalisasi data untuk menstandarkan skala, pengelompokan kategori agar lebih mudah dianalisis, serta pembuatan variabel baru yang relevan dengan kebutuhan analisis. Dengan menjalankan proses ETL secara optimal, kualitas data menjadi lebih terjamin, dan analisis yang dilakukan pada tahap *Modeling* akan menghasilkan *output* yang lebih valid dan bermanfaat. Oleh karena itu, tahap ini tidak hanya bersifat teknis, tetapi juga strategis, karena berperan penting dalam memastikan integritas data dan keberhasilan keseluruhan proyek analitik di lingkungan rumah sakit.

3.2.4 Menganalisis Data Menggunakan Teknik Statistik dan Alat Analitik

Dalam analisis data layanan kesehatan, teknik statistik dan alat analitik digunakan untuk menggali wawasan yang mendalam yang mendukung pengambilan keputusan berbasis data yang lebih tepat. Proses ini merupakan bagian dari tahap *Modeling* dalam CRISP-DM, yang

bertujuan untuk membangun model prediktif atau deskriptif yang dapat membantu dalam perencanaan dan pengambilan keputusan. Proses ini dimulai dengan identifikasi pola-pola tren berdasarkan data. Selanjutnya, *data mining* diterapkan untuk memprediksi kebutuhan stok obat, mengidentifikasi jenis obat yang paling banyak dibutuhkan pada waktu tertentu, dan memprediksi potensi kekurangan stok obat di masa depan.. Dengan menggunakan teknik analitik yang tepat, informasi yang diperoleh tidak hanya memberikan wawasan yang lebih tajam mengenai dinamika layanan kesehatan, tetapi juga membantu dalam perencanaan strategis yang lebih efektif, seperti merencanakan pengadaan obat-obatan, merumuskan kebijakan distribusi yang lebih efisien, serta meningkatkan kualitas perawatan pasien. Selain itu, optimalisasi sumber daya, baik itu tenaga medis, obat-obatan, maupun fasilitas rumah sakit, menjadi lebih mudah dilakukan, yang pada akhirnya berdampak pada peningkatan efisiensi operasional secara keseluruhan. Penerapan teknik ini juga mendukung pengambilan keputusan yang lebih cepat dan berbasis data, yang membantu rumah sakit dalam menanggapi perubahan kebutuhan pasien dengan lebih responsif dan proaktif.

1. Mengidentifikasi Pola Tren Layanan Kesehatan Berdasarkan Data

Langkah awal dalam analisis data adalah melakukan eksplorasi awal terhadap dataset yang tersedia. Seperti yang ditampilkan pada Gambar 3.9, proses ini dimulai dengan memuat data penjualan obat dari file Excel bernama LaporanPenjualan_Pivot.xlsx menggunakan bahasa pemrograman Python. Data ini disajikan dalam bentuk tabel pivot, dengan kolom-kolom yang merepresentasikan jumlah penjualan obat per bulan sepanjang tahun 2024, dan baris-baris yang menunjukkan kode item obat. Melalui tampilan awal data ini, kita dapat langsung mengamati pola-pola tertentu, misalnya obat dengan kode OBT00004 yang menunjukkan lonjakan penjualan signifikan pada bulan Februari dan Mei, sementara obat lain seperti OBT00003 memiliki fluktuasi yang lebih tidak stabil.

```

import pandas as pd
import numpy as np

df = pd.read_excel('LaporanPenjualan_Pivot.xlsx')
print(df.head())

```

Kode Item	January 2024	February 2024	March 2024	April 2024	May 2024
0 OBT00001	0	0	0	0	0
1 OBT00004	35	131	10	15	124
2 OBT00006	120	0	38	95	40
3 OBT00008	30	5	-2	14	37
4 OBT00010	1	0	0	4	2

	June 2024	July 2024	August 2024	September 2024	October 2024
0	0	0	0	0	0
1	61	24	21	181	171
2	0	63	111	90	180
3	45	53	40	15	0
4	0	0	2	0	0

	November 2024	December 2024
0	14	0
1	413	295
2	160	50
3	30	0
4	0	0

Gambar 3.8 Kode Identifikasi Pola Tren

Identifikasi pola seperti ini sangat penting karena dapat memberikan wawasan tentang bagaimana perilaku permintaan terhadap obat tertentu berubah seiring waktu. Misalnya, permintaan tinggi pada bulan tertentu bisa jadi karena musim penyakit tertentu atau perubahan kebijakan layanan. Dari sini, pengelola fasilitas layanan kesehatan dapat mengoptimalkan penyediaan sumber daya, menyesuaikan anggaran, dan menghindari kekurangan obat. Selain itu, data ini juga membantu dalam mendeteksi potensi masalah lebih awal, misalnya jika ada obat yang penjualannya turun drastis, bisa jadi karena pasokannya terhambat atau karena efektivitasnya dipertanyakan oleh dokter.

Tren penggunaan obat juga dapat menunjukkan respons masyarakat terhadap wabah atau kondisi kesehatan tertentu yang muncul secara tiba-tiba. Misalnya, lonjakan mendadak dalam penjualan obat penurun demam dan antibiotik tertentu bisa menjadi indikasi awal adanya penyebaran penyakit infeksi di masyarakat, bahkan sebelum data klinis dari rumah sakit dikumpulkan secara resmi. Oleh karena itu, pemantauan pola penjualan obat secara *real-time* dapat menjadi sumber data alternatif yang sangat berguna dalam sistem pemantauan epidemiologis dan pengambilan keputusan cepat dalam kondisi darurat.

Selain itu, analisis tren penggunaan obat juga bermanfaat dalam mengevaluasi efektivitas program kesehatan yang telah dijalankan. Misalnya, jika setelah kampanye besar-besaran tentang pengendalian hipertensi terlihat adanya peningkatan pembelian obat antihipertensi generik, maka ini dapat menjadi indikator bahwa kesadaran masyarakat terhadap pengobatan rutin mulai meningkat. Sebaliknya, jika tren penjualan obat justru menurun, bisa jadi perlu dilakukan evaluasi terhadap strategi komunikasi atau distribusi obat yang digunakan. Dengan kata lain, tren obat dapat mencerminkan keberhasilan atau kegagalan intervensi kesehatan masyarakat.

Lebih jauh, pemahaman terhadap tren penggunaan obat juga membantu dalam merancang kebijakan pengadaan yang efisien dan berbasis bukti. Obat-obat yang menunjukkan tren permintaan konsisten dari bulan ke bulan dapat diprioritaskan dalam pengadaan jangka panjang, sedangkan obat dengan permintaan yang sangat fluktuatif perlu dikelola dengan strategi yang lebih adaptif. Pendekatan ini tidak hanya membantu mengurangi pemborosan anggaran akibat stok obat kedaluwarsa, tetapi juga memastikan ketersediaan obat yang tepat, pada waktu yang tepat, dan dalam jumlah yang sesuai dengan kebutuhan nyata di lapangan.

2. Melakukan Prediksi Stok Obat Menggunakan *Data mining*

Prediksi stok obat dilakukan untuk memastikan ketersediaan obat yang optimal berdasarkan pola permintaan yang telah dianalisis. Dengan menggunakan teknik *data mining*, model prediktif dikembangkan untuk memperkirakan jumlah stok yang dibutuhkan dalam periode tertentu. Analisis ini didasarkan pada data historis penjualan obat, termasuk total item yang terjual, frekuensi pembelian, serta tren permintaan berdasarkan bulan. Dengan pendekatan ini, stok obat dapat dikelola secara lebih efisien, mengurangi risiko kekurangan atau kelebihan persediaan, serta meningkatkan efektivitas perencanaan farmasi di fasilitas layanan kesehatan. Setelah memahami tren masa lalu, langkah berikutnya adalah

memanfaatkan data historis untuk membuat prediksi yang berguna ke depan. Di sinilah teknik *data mining* digunakan. Pada Gambar 3.10, ditunjukkan bagaimana model prediktif dibuat menggunakan algoritma *XGBoost* dan *Random Forest*, sebuah metode *machine learning* yang cukup andal untuk memproyeksikan nilai berdasarkan data masa lalu. Model ini dilatih dengan data penjualan obat dari bulan Januari hingga Desember 2024, dengan tujuan utama untuk memprediksi kebutuhan stok di tahun 2025.

Gambar 3.10 ini menampilkan tahap awal dalam proses analisis data penjualan obat, yaitu proses pengunggahan file Excel bernama *LaporanPenjualan_Pivot.xlsx* ke dalam lingkungan kerja Google Colab. File ini berisi data penjualan obat dalam bentuk pivot, yaitu setiap kolom mewakili bulan dan setiap baris mewakili kode item. File yang telah diunggah disimpan ulang secara otomatis agar dapat digunakan dalam pemrosesan selanjutnya. Tahap ini merupakan langkah penting untuk memastikan bahwa data historis tersedia dalam format yang bisa dibaca oleh sistem pemodelan prediktif.

```

[1] from google.colab import files
    uploaded = files.upload()

Choose Files | LaporanPe...n_Pivot.xlsx
• LaporanPenjualan_Pivot.xlsx(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 52015 bytes, last modified: 4/13/2025 - 100% done
Saving LaporanPenjualan_Pivot.xlsx to LaporanPenjualan_Pivot.xlsx

[2] # 1. Import Library
import pandas as pd
import numpy as np
import joblib
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
import matplotlib.pyplot as plt
from xgboost import XGBRegressor

# 2. Baca Data & Ubah Format
df = pd.read_excel('LaporanPenjualan_Pivot.xlsx')

df_long = (
df.melt(id_vars="Kode Item", var_name="Bulan", value_name="Penjualan")
.assign(
    Bulan=lambda d: pd.to_datetime(d["Bulan"], format="%B %Y"),
    Penjualan=lambda d: d["Penjualan"].fillna(0).clip(lower=0)
)
.sort_values(["Kode Item", "Bulan"])
.reset_index(drop=True)
)

```

Gambar 3.9 Proses Awal Analisis Data

Langkah selanjutnya adalah melakukan *import* berbagai *library* yang diperlukan untuk proses analisis dan pemodelan. *Library* seperti *pandas* dan *numpy* digunakan untuk manipulasi data, sementara *sklearn* menyediakan alat untuk pelatihan model prediktif, preprocessing data, dan evaluasi model. Selain itu, *XGBoost* juga diimpor karena digunakan sebagai salah satu algoritma utama dalam proses prediksi. Dengan mengimpor *library* ini, seluruh fungsi yang diperlukan dalam pipeline *data mining* seperti encoding, pembagian data, dan pemodelan telah disiapkan.

Data yang awalnya dalam format *pivot* kemudian diubah ke format *long* (panjang) menggunakan fungsi *melt()* dari *pandas*. Transformasi ini penting agar setiap baris mencerminkan satu observasi unik yakni kombinasi antara Kode Item dan Bulan dengan nilai Penjualan yang sesuai. Selain itu, kolom Bulan dikonversi menjadi format tanggal, dan nilai penjualan yang kosong diisi dengan nol untuk menjaga integritas data. Nilai negatif juga dihindari dengan penggunaan fungsi *clip(lower=0)*, sehingga hanya nilai penjualan valid yang dianalisis. Data diurutkan berdasarkan Kode Item dan Bulan untuk memastikan urutan kronologis yang sesuai bagi pemodelan berbasis waktu.

Gambar 3.11 memperlihatkan proses penting dalam tahap *feature engineering*, yaitu pembuatan fitur tambahan yang dapat meningkatkan akurasi model prediksi. Dalam proses ini, sejumlah fitur baru dikembangkan berdasarkan data penjualan sebelumnya. Fitur-fitur seperti *Penjualan_Lag_1*, *Penjualan_Lag_2*, dan *Penjualan_Lag_3* merepresentasikan penjualan item pada satu, dua, dan tiga bulan sebelumnya. Informasi ini penting karena memungkinkan model memahami pola historis dan menjadikannya acuan untuk memprediksi nilai di bulan berikutnya.

```

[27] # 3. Tambahkan Fitur Tambahan
df_long["Penjualan_Lag_1"] = df_long.groupby("Kode Item")["Penjualan"].shift(1).fillna(0)
df_long["Penjualan_Lag_2"] = df_long.groupby("Kode Item")["Penjualan"].shift(2).fillna(0)
df_long["Penjualan_Lag_3"] = df_long.groupby("Kode Item")["Penjualan"].shift(3).fillna(0)
df_long["Sales_Diff_1"] = df_long.groupby("Kode Item")["Penjualan"].diff(1).fillna(0)
df_long["Rolling_Mean_3"] = df_long.groupby("Kode Item")["Penjualan"].transform(lambda x: x.rolling(3, min_periods=1).mean())
df_long["Bulan_Num"] = df_long["Bulan"].dt.month
df_long["Tahun"] = df_long["Bulan"].dt.year

[5] # 4. Split Data Train / Validasi (Validasi: Okt-Des 2024)
mask_val = df_long["Bulan"] >= "2024-10-01"
train_df = df_long.loc[~mask_val].copy()
val_df = df_long.loc[mask_val].copy()

[6] # 5. Label Encoding 'Kode Item'
le = LabelEncoder()
train_df["KodeItem_enc"] = le.fit_transform(train_df["Kode Item"])
val_df["KodeItem_enc"] = le.transform(val_df["Kode Item"])

# 6. Siapkan Fitur Model
FEATURES = [
    "KodeItem_enc", "Bulan_Num",
    "Penjualan_Lag_1", "Penjualan_Lag_2", "Penjualan_Lag_3",
    "Sales_Diff_1", "Rolling_Mean_3"
]

X_train, y_train = train_df[FEATURES], train_df["Penjualan"]
X_val, y_val = val_df[FEATURES], val_df["Penjualan"]

```

Gambar 3.10 Proses *Feature Engineering*

Selain itu, fitur `Sales_Diff_1` dibuat untuk menangkap perubahan penjualan dari bulan ke bulan. Fitur ini menghitung selisih antara penjualan saat ini dan penjualan bulan sebelumnya, sehingga dapat memberikan sinyal kepada model apakah terjadi peningkatan atau penurunan permintaan. Kemudian, fitur `Rolling_Mean_3` digunakan untuk menghitung rata-rata penjualan selama tiga bulan terakhir. Ini berguna dalam mengidentifikasi tren jangka pendek dan memberikan gambaran stabilitas penjualan suatu item.

Model juga dibekali dengan informasi waktu, seperti bulan dan tahun, yang diubah menjadi format numerik melalui fitur `Bulan_Num` dan `Tahun`. Transformasi ini penting karena beberapa pola penjualan mungkin dipengaruhi oleh musim, liburan, atau siklus tahunan. Dengan mengonversi waktu ke dalam bentuk angka, model dapat lebih mudah mengenali keterkaitan tersebut dalam data. Fitur-fitur tambahan yang telah dibuat ini secara keseluruhan membantu model dalam memahami struktur dan pola data penjualan yang tidak terlihat secara langsung dari data mentah. Berikut merupakan deskripsi *feature engineering* dalam bentuk tabel 3.5.

Tabel 3.5 Deskripsi Feature Engineering

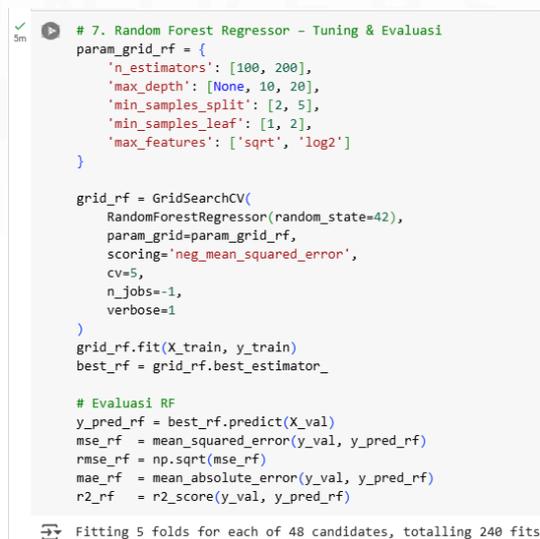
Nama Fitur	Metode Pembentukan	Deskripsi
Penjualan_Lag_1	Shift data penjualan 1 bulan ke belakang	Jumlah penjualan item pada 1 bulan sebelumnya.
Penjualan_Lag_2	Shift data penjualan 2 bulan ke belakang	Jumlah penjualan item pada 2 bulan sebelumnya.
Penjualan_Lag_3	Shift data penjualan 3 bulan ke belakang	Jumlah penjualan item pada 3 bulan sebelumnya.
Sales_Diff_1	Selisih (Penjualan saat ini-Penjualan_Lag_1)	Mengukur perubahan penjualan dari bulan sebelumnya.
Rolling_Mean_3	Rata-rata bergerak 3 bulan terakhir	Menunjukkan tren penjualan jangka pendek berdasarkan rata-rata 3 bulan.
Bulan_Num	Konversi nama bulan ke angka (Januari = 1,dst)	Representasi numerik bulan untuk analisis musiman.
Tahun	Ekstrasi tahun dari kolom waktu	Menyediakan konteks tahunan dalam analisis pola data.

Langkah berikutnya adalah memisahkan data menjadi dua bagian: data latih (*train_df*) dan data validasi (*val_df*). Data sebelum Oktober 2024 digunakan sebagai data latih, sementara data dari bulan Oktober hingga Desember 2024 digunakan sebagai data validasi. Tujuan dari pembagian ini adalah agar model dapat dilatih menggunakan data historis, kemudian diuji pada data terbaru yang mendekati periode prediksi. Dengan cara ini, hasil prediksi untuk tahun 2025 nantinya diharapkan menjadi lebih akurat dan andal.

Sebelum model dilatih, kolom kategorikal seperti Kode Item yang berisi nama produk perlu diubah ke dalam format numerik agar bisa dibaca oleh algoritma *machine learning*. Proses ini dikenal sebagai label encoding, di mana setiap kategori diubah menjadi representasi angka unik. Setelah proses encoding selesai, dipilih sejumlah fitur penting yang akan digunakan dalam pelatihan model, seperti fitur lag (nilai historis), fitur waktu (seperti bulan atau hari), serta rolling mean (rata-rata bergerak). Dataset kemudian dipisahkan menjadi fitur (*X_train*, *X_val*) dan target (*y_train*, *y_val*) yang siap digunakan dalam pelatihan dan evaluasi model prediksi stok obat.

Setelah proses label encoding selesai, daftar fitur input disusun berdasarkan kolom-kolom hasil *feature engineering* sebelumnya. Data input (X_{train} , X_{val}) dan target (y_{train} , y_{val}) untuk model disiapkan dengan memilih kolom-kolom yang relevan sebagai fitur serta mengatur kolom target yang berisi nilai penjualan obat. Langkah ini memastikan bahwa model hanya menerima data yang penting dan sesuai untuk pelatihan dan evaluasi, sehingga prediksi stok menjadi lebih optimal. Proses pemilihan fitur yang tepat juga membantu mengurangi kompleksitas model dan mencegah *overfitting*. Dengan demikian, model diharapkan dapat menghasilkan prediksi yang lebih stabil dan akurat pada data validasi dan data baru di masa depan.

Gambar 3.12 memperlihatkan proses pelatihan dan evaluasi model *Random Forest* Regressor menggunakan teknik *GridSearchCV*. Teknik ini digunakan untuk mencari kombinasi *hyperparameter* terbaik dari beberapa parameter penting seperti *n_estimators*, *max_depth*, *min_samples_split*, *min_samples_leaf*, dan *max_features*. Masing-masing parameter memiliki nilai-nilai yang telah ditentukan dalam sebuah grid, yang kemudian dievaluasi satu per satu oleh algoritma untuk menemukan konfigurasi paling optimal.



```
# 7. Random Forest Regressor - Tuning & Evaluasi
param_grid_rf = {
    'n_estimators': [100, 200],
    'max_depth': [None, 10, 20],
    'min_samples_split': [2, 5],
    'min_samples_leaf': [1, 2],
    'max_features': ['sqrt', 'log2']
}

grid_rf = GridSearchCV(
    RandomForestRegressor(random_state=42),
    param_grid=param_grid_rf,
    scoring='neg_mean_squared_error',
    cv=5,
    n_jobs=-1,
    verbose=1
)
grid_rf.fit(X_train, y_train)
best_rf = grid_rf.best_estimator_

# Evaluasi RF
y_pred_rf = best_rf.predict(X_val)
mse_rf = mean_squared_error(y_val, y_pred_rf)
rmse_rf = np.sqrt(mse_rf)
mae_rf = mean_absolute_error(y_val, y_pred_rf)
r2_rf = r2_score(y_val, y_pred_rf)
```

Fitting 5 folds for each of 48 candidates, totalling 240 fits

Gambar 3.11 Proses Tuning dan Evaluasi Model *Random Forest* Regressor

Proses pencarian parameter terbaik dilakukan dengan metode validasi silang (cross-validation) sebanyak lima lipatan ($cv=5$). Hal ini berarti data pelatihan dibagi menjadi lima bagian, di mana model akan dilatih pada empat bagian dan divalidasi pada satu bagian secara bergantian. Evaluasi dilakukan menggunakan metrik `neg_mean_squared_error`, yaitu versi negatif dari mean squared error (MSE). Penggunaan nilai negatif bertujuan agar skor evaluasi tetap bisa dimaksimalkan oleh `GridSearchCV`, karena secara default fungsi ini mencari nilai tertinggi.

Setelah semua kombinasi diuji, `GridSearchCV` akan memilih model dengan performa terbaik yang kemudian disimpan dalam variabel `best_rf`. Model ini digunakan untuk memprediksi nilai pada data validasi (`X_val`). Evaluasi hasil prediksi dilakukan dengan menghitung beberapa metrik performa seperti MSE (mean squared error), RMSE (root mean squared error), MAE (mean absolute error), dan R^2 Score. RMSE memberikan gambaran kesalahan dalam satuan aslinya, sedangkan MAE menunjukkan rata-rata kesalahan absolut, dan R^2 Score mengukur proporsi variansi yang dapat dijelaskan oleh model.

Gambar 3.13 memperlihatkan proses *tuning* dan evaluasi model `XGBoost` Regressor, yaitu salah satu algoritma ensemble berbasis pohon yang sangat populer dalam berbagai kompetisi prediksi karena kecepatan dan akurasi. Proses *tuning* dilakukan dengan menggunakan `GridSearchCV`, sebuah metode sistematis untuk mencari kombinasi hyperparameter terbaik. Parameter yang diuji mencakup `n_estimators`, `learning_rate`, `max_depth`, `subsample`, `colsample_bytree`, dan `reg_lambda`. Seluruh parameter ini berpengaruh terhadap kompleksitas dan performa model secara keseluruhan.

```

# 8. XGBoost Regressor - Tuning & Evaluasi
param_grid_xgb = {
    "n_estimators": [300, 500],
    "learning_rate": [0.05, 0.1],
    "max_depth": [4, 6, 8],
    "subsample": [0.8, 1.0],
    "colsample_bytree": [0.8, 1.0],
    "reg_lambda": [1.0, 1.5],
}

grid_xgb = GridSearchCV(
    XGBRegressor(
        objective="reg:squarederror",
        random_state=42,
        tree_method="hist"
    ),
    param_grid=param_grid_xgb,
    scoring="neg_mean_squared_error",
    cv=5,
    n_jobs=-1,
    verbose=1,
)

grid_xgb.fit(X_train, y_train)
best_xgb = grid_xgb.best_estimator_

# Evaluasi XGB
y_pred_xgb = best_xgb.predict(X_val)
mse_xgb = mean_squared_error(y_val, y_pred_xgb)
rmse_xgb = np.sqrt(mse_xgb)
mae_xgb = mean_absolute_error(y_val, y_pred_xgb)
r2_xgb = r2_score(y_val, y_pred_xgb)

```

Fitting 5 folds for each of 96 candidates, totalling 480 fits

Gambar 3.12 Proses Tuning dan Evaluasi Model *XGBoost* Regressor

Untuk mencari konfigurasi terbaik, GridSearchCV melakukan proses validasi silang (cross-validation) sebanyak lima lipatan ($cv=5$). Artinya, data pelatihan dibagi menjadi lima bagian dan model dilatih serta diuji secara bergantian pada masing-masing bagian. Evaluasi model dilakukan menggunakan metrik `neg_mean_squared_error`, yaitu versi negatif dari rata-rata kuadrat kesalahan prediksi. Metrik ini digunakan karena GridSearchCV secara default mengutamakan skor maksimum, sehingga nilai kesalahan dikonversi menjadi negatif.

Setelah seluruh kombinasi parameter diuji, GridSearchCV memilih model *XGBoost* dengan performa terbaik dan menyimpannya dalam variabel `best_xgb`. Model terbaik ini kemudian digunakan untuk melakukan prediksi terhadap data validasi (`X_val`). Hasil prediksi dibandingkan dengan nilai aktual (`y_val`) menggunakan beberapa metrik evaluasi seperti MSE, RMSE, MAE, dan R^2 Score. Metrik-metrik ini memberikan gambaran menyeluruh tentang akurasi dan efektivitas model dalam merepresentasikan data.

Melalui proses yang digambarkan pada Gambar 3.13, model *XGBoost* disiapkan secara optimal untuk mendukung prediksi stok obat dengan tingkat

akurasi tinggi. *Tuning hyperparameter* secara sistematis membantu mencegah *overfitting* dan meningkatkan kemampuan generalisasi model terhadap data baru. Kombinasi antara kekuatan algoritma *XGBoost* dan evaluasi mendalam menjadikan pendekatan ini sangat efektif dalam menangkap pola kompleks dari data historis. Dengan demikian, model yang dihasilkan dapat memberikan prediksi yang stabil, akurat, dan siap digunakan dalam pengambilan keputusan operasional.

Gambar 3.14 memperlihatkan proses perbandingan kinerja dua model prediksi, yaitu *Random Forest* dan *XGBoost*, berdasarkan empat metrik evaluasi utama: MSE, RMSE, MAE, dan R^2 Score. Perbandingan ini dilakukan dengan menggabungkan hasil evaluasi kedua model ke dalam sebuah DataFrame pandas yang disusun secara sistematis. Masing-masing metrik menggambarkan seberapa akurat dan stabil model dalam melakukan prediksi terhadap data validasi. Proses ini penting untuk menentukan model mana yang paling optimal digunakan dalam prediksi stok obat.

```

# 9. Bandingkan Model
comparison = pd.DataFrame({
    "Metric": ["MSE", "RMSE", "MAE", "R²"],
    "Random Forest": [mse_rf, rmse_rf, mae_rf, r2_rf],
    "XGBoost": [mse_xgb, rmse_xgb, mae_xgb, r2_xgb],
})

print("\n Perbandingan Model:")
print(comparison)

```

Perbandingan Model:			
Metric	Random Forest	XGBoost	
0	MSE	735.799531	652.555359
1	RMSE	27.125625	25.545163
2	MAE	5.141064	3.404639
3	R ²	0.937630	0.944686

Gambar 3.13 Perbandingan Model *XGBoost* dan *Random Forest*

Dari hasil perbandingan, model *XGBoost* menunjukkan performa yang lebih baik dibandingkan *Random Forest* pada semua metrik evaluasi. Nilai MSE pada *XGBoost* sebesar 652.56 lebih rendah dibandingkan *Random Forest* yang mencapai 735.80, menunjukkan bahwa rata-rata kuadrat kesalahannya lebih kecil. Demikian juga dengan RMSE, yang merepresentasikan akar dari MSE, *XGBoost* menghasilkan nilai 25.54, lebih kecil daripada 27.13 milik *Random Forest*. Nilai

ini menunjukkan bahwa prediksi *XGBoost* lebih dekat dengan nilai aktual dibandingkan *Random Forest*.

Selain itu, nilai MAE pada model *XGBoost* juga lebih rendah, yaitu 3.40 dibandingkan 5.14 pada *Random Forest*, yang berarti bahwa rata-rata kesalahan absolut model *XGBoost* lebih kecil. R^2 Score sebagai metrik koefisien determinasi menunjukkan nilai 0.9447 untuk *XGBoost* dan 0.9376 untuk *Random Forest*. Nilai ini mengindikasikan bahwa *XGBoost* mampu menjelaskan sekitar 94,47% variasi data, sedikit lebih baik dari *Random Forest*. Perbedaan ini, walaupun tidak sangat besar, cukup berarti dalam konteks prediksi stok yang memerlukan tingkat presisi tinggi.

Keseluruhan proses pada Gambar 3.14 menegaskan bahwa model *XGBoost* lebih unggul dalam memodelkan pola data historis penjualan obat dibandingkan *Random Forest*. Evaluasi menyeluruh menggunakan berbagai metrik memberikan dasar yang kuat dalam memilih model terbaik untuk diterapkan pada data baru. Dengan performa prediksi yang lebih akurat dan kestabilan model yang lebih baik, *XGBoost* menjadi pilihan utama dalam sistem prediksi stok. Hasil ini juga memperlihatkan pentingnya melakukan proses komparatif setelah pelatihan model untuk memastikan keputusan berbasis data yang tepat.

Gambar 3.15 menampilkan proses lengkap untuk melakukan prediksi penjualan obat selama tahun 2025. Pemilihan model terbaik dilakukan berdasarkan nilai RMSE terkecil antara dua kandidat, yaitu *Random Forest* dan *XGBoost*. Model dengan kinerja terbaik disimpan dalam variabel `chosen_model` dan digunakan untuk menghasilkan prediksi ke depan. Dengan pendekatan ini, hanya model yang memiliki performa terbaik pada data validasi yang digunakan dalam proses prediksi tahunan.

```

import pandas as pd

# 1. Pilih model terbaik
chosen_model = best_xgb if rmse_xgb < rmse_rf else best_rf
rmse_chosen = rmse_xgb if chosen_model == best_xgb else rmse_rf

# 2. Prediksi untuk data validasi
val_df["Prediksi"] = chosen_model.predict(X_val)

# 3. Buat data awal untuk Januari 2025
future_df = val_df.groupby("Kode Item").tail(1).copy()
future_df["Bulan"] = pd.to_datetime("2025-01-01")
future_df["Bulan_Num"] = 1

# 4. Simpan hasil prediksi semua bulan
all_predictions = []

# 5. Loop prediksi Januari-Desember 2025
for month in range(1, 13):
    current_df = future_df.copy()
    current_df["Bulan"] = pd.to_datetime(f"2025-{month:02d}-01")
    current_df["Bulan_Num"] = month

    # Update fitur lag
    current_df["Penjualan_Lag_1"] = future_df["Prediksi"]
    current_df["Penjualan_Lag_2"] = future_df.get("Penjualan_Lag_1", 0)
    current_df["Penjualan_Lag_3"] = future_df.get("Penjualan_Lag_2", 0)

    # Fitur turunan
    current_df["Sales_Diff_1"] = current_df["Penjualan_Lag_1"] - current_df["Penjualan_Lag_2"]
    current_df["Rolling_Mean_3"] = current_df[["Penjualan_Lag_1", "Penjualan_Lag_2", "Penjualan_Lag_3"]].mean(axis=1)
    current_df["KodeItem_enc"] = le.transform(current_df["Kode Item"])

    # Prediksi
    current_df["Prediksi"] = chosen_model.predict(current_df[FEATURES])

```

Gambar 3.14 Prediksi Menggunakan Model Terbaik

Langkah selanjutnya adalah membuat prediksi awal pada data validasi dan menyimpannya dalam kolom baru bernama “Prediksi”. Data terakhir dari penjualan digunakan sebagai basis awal untuk membuat DataFrame `future_df` yang mewakili kondisi bulan Januari 2025. Tanggal disesuaikan dengan awal tahun 2025 dan informasi numerik bulan ditambahkan. Hal ini penting agar format waktu dapat dibaca dengan benar oleh model selama proses prediksi berlangsung.

Proses prediksi dilakukan secara iteratif untuk bulan Januari hingga Desember 2025 dengan menggunakan perulangan sebanyak 12 kali. Dalam setiap iterasi, salinan data dasar diperbarui agar mencerminkan bulan yang sedang diprediksi. Informasi waktu seperti tanggal dan nomor bulan diperbarui agar relevan dengan iterasi tersebut. Dengan metode ini, prediksi disusun secara berjenjang mengikuti urutan waktu secara akurat.

Di setiap langkah iterasi, fitur-fitur historis seperti `Penjualan_Lag_1`, `Penjualan_Lag_2`, dan `Penjualan_Lag_3` diperbarui dari hasil prediksi bulan sebelumnya. Selain itu, fitur turunan seperti `Sales_Diff_1` dan `Rolling_Mean_3` dihitung ulang untuk mencerminkan dinamika penjualan terbaru. Kolom kategori seperti `Kode Item` juga diencoding ulang agar tetap kompatibel dengan input model. Seluruh pembaruan ini dilakukan secara otomatis di setiap bulan untuk menjaga konsistensi data input.

Prediksi bulanan kemudian dilakukan dengan memanggil metode `.predict()` pada model terbaik menggunakan fitur yang telah diperbarui. Hasil prediksi disimpan dalam kolom “Prediksi” untuk digunakan dalam iterasi bulan berikutnya. Setiap hasil prediksi juga ditambahkan ke daftar `all_predictions` yang menyimpan prediksi dari Januari hingga Desember. Strategi ini merupakan pendekatan recursive forecasting yang umum digunakan dalam analisis deret waktu untuk memastikan kesinambungan antar bulan.

Pada bagian awal Gambar 3.16, proses dimulai dengan memastikan bahwa setiap nilai prediksi memiliki batas minimal 1. Hal ini penting karena stok obat tidak mungkin bernilai nol atau negatif; maka dari itu, fungsi `apply(lambda x: max(round(x), 1))` digunakan untuk memastikan nilai Prediksi selalu minimal satu. Selanjutnya, ditentukan nilai `Prediksi_Min` dan `Prediksi_Max`, yaitu nilai prediksi yang dikurangi dan ditambah dengan error model (`nmse_chosen`) untuk membentuk rentang estimasi. Nilai ini memberikan margin of error yang membantu pihak manajemen rumah sakit dalam mempertimbangkan ketidakpastian prediksi saat menentukan jumlah pengadaan stok obat.

```

# Pastikan Prediksi minimal 1
current_df["Prediksi"] = current_df["Prediksi"].apply(lambda x: max(round(x), 1))

# Prediksi Min dan Max
current_df["Prediksi_Min"] = current_df["Prediksi"] - round(nmse_chosen)
current_df["Prediksi_Min"] = current_df["Prediksi_Min"].apply(lambda x: max(x, 1))
current_df["Prediksi_Max"] = current_df["Prediksi"] + round(nmse_chosen)

# Label bulan
current_df["Bulan_Prediksi"] = current_df["Bulan"].dt.strftime("%Y-%m")

# Simpan hasil
all_predictions.append(current_df[["Kode Item", "Bulan_Prediksi", "Prediksi", "Prediksi_Min", "Prediksi_Max"]])

# Update untuk bulan berikutnya
future_df = current_df.copy()

# 6. Gabungkan seluruh hasil prediksi
result_df = pd.concat(all_predictions, ignore_index=True)

# 7. Pastikan pembulatan dan nilai minimal
result_df["Prediksi"] = result_df["Prediksi"].round(0).astype(int)
result_df["Prediksi_Min"] = result_df["Prediksi_Min"].round(0).astype(int)
result_df["Prediksi_Max"] = result_df["Prediksi_Max"].round(0).astype(int)

# 8. Label bulan untuk pivot
result_df["Bulan_Label"] = pd.to_datetime(result_df["Bulan_Prediksi"]).dt.strftime("%B-%Y")

# 9. Buat pivot table
pivot_df = result_df.pivot(index="Kode Item", columns="Bulan_Label", values=["Prediksi", "Prediksi_Min", "Prediksi_Max"])

# 10. Susun kolom berdasarkan urutan bulan
ordered_months = pd.date_range("2025-01-01", "2025-12-01", freq="MS").strftime("%B-%Y")
ordered_columns = [(jenis, bulan) for bulan in ordered_months for jenis in ["Prediksi", "Prediksi_Min", "Prediksi_Max"]]
pivot_df = pivot_df[ordered_columns]

# 11. Ratakan kolom MultiIndex
pivot_df.columns = ["{jenis}_{bulan}" for jenis, bulan in pivot_df.columns]

# 12. Simpan ke Excel
output_file = "Prediksi_Stok_Obat_2025.xlsx"
pivot_df.to_excel(output_file, index=True)

print("\n✅ File prediksi lengkap (Jan-Des 2025) disimpan sebagai:")
print(f"📁 {output_file}")

```

File prediksi lengkap (Jan-Des 2025) disimpan sebagai:
 Prediksi Stok Obat 2025.xlsx

Gambar 3.15 Prediksi Stok Obat Menggunakan *XGBoost*

Setelah nilai prediksi dan rentangnya dihitung, ditambahkan kolom Bulan_Prediksi yang dihasilkan dari kolom Bulan dengan format "%Y-%m". Format ini membuat data prediksi lebih mudah dikelompokkan dan dianalisis secara bulanan. Dataframe hasil prediksi untuk bulan tersebut kemudian disimpan ke dalam list bernama all_predictions dengan menyertakan kolom Kode Item, Bulan_Prediksi, dan ketiga jenis prediksi. Setelah itu, dataframe future_df diperbarui untuk digunakan dalam iterasi bulan selanjutnya, sehingga proses prediksi bisa dilakukan secara berulang dari bulan ke bulan.

Selanjutnya, setelah prediksi untuk seluruh bulan selesai, semua dataframe yang tersimpan dalam all_predictions digabungkan menggunakan fungsi pd.concat() menjadi dataframe tunggal result_df. Hal ini menyatukan semua hasil prediksi dari Januari hingga Desember 2025. Kemudian, seluruh nilai prediksi, termasuk Prediksi, Prediksi_Min, dan Prediksi_Max, dibulatkan ke angka bulat dan dikonversi menjadi tipe data int. Hal ini dilakukan untuk menyajikan angka yang bersih dan siap digunakan dalam dokumen laporan atau sistem informasi manajemen rumah sakit.

Untuk keperluan visualisasi dan penyusunan tabel pivot, dibuat kolom baru Bulan_Label yang merupakan hasil konversi kolom Bulan_Prediksi ke format string seperti "Jan-2025", "Feb-2025", dan seterusnya. Langkah ini mempermudah pengguna dalam membaca dan menelusuri data bulanan secara kronologis. Kolom Bulan_Label akan digunakan sebagai label kolom dalam pivot table, karena lebih informatif dibandingkan format numerik. Dengan ini, data prediksi akan lebih mudah dipahami oleh pengguna non-teknis seperti staf logistik atau farmasi.

Pada tahap berikutnya, pivot table dibuat dengan menggunakan Kode Item sebagai index, Bulan_Label sebagai kolom, dan isi cell berupa nilai prediksi. Pivot table ini menyusun data dalam format dua dimensi agar mudah digunakan sebagai laporan stok. Kemudian disusun urutan kolom berdasarkan bulan dari Januari hingga Desember 2025 menggunakan pd.date_range() dan format "%b-%Y" . Proses ini memastikan bahwa kolom-kolom prediksi berada dalam urutan

kronologis yang benar dan konsisten untuk seluruh jenis prediksi (Prediksi, Prediksi_Min, dan Prediksi_Max).

Kolom pada pivot table awalnya merupakan kolom bertingkat (MultiIndex), sehingga perlu diratakan agar lebih sederhana dan siap digunakan di Excel. Hasil pelurusan kolom akan menghasilkan nama seperti Prediksi_Jan-2025, Prediksi_Min_Jan-2025, dan seterusnya, yang lebih mudah dipahami. Setelah pivot table tersusun dengan rapi, hasil akhirnya disimpan ke file Excel bernama Prediksi_Stok_Obat_2025.xlsx menggunakan `to_excel()`. File ini menjadi output akhir dari seluruh proses prediksi dan dapat digunakan untuk perencanaan pengadaan obat di Rumah Sakit Mentari selama periode Januari hingga Desember 2025.

Gambar 3.17 menunjukkan proses penyimpanan model *machine learning* dan encoder label ke dalam file menggunakan pustaka `joblib`. Pada bagian ini, tiga objek penting disimpan: model *Random Forest* terbaik (`best_rf`), model *XGBoost* terbaik (`best_xgb`), serta objek `LabelEncoder` (`le`) yang digunakan untuk mengubah data kategorikal menjadi numerik selama proses pelatihan. Penyimpanan ini penting agar model tidak perlu dilatih ulang setiap kali digunakan, sehingga dapat langsung digunakan untuk prediksi pada data baru. File hasil penyimpanan masing-masing diberi nama `model_rf.pkl`, `model_xgb.pkl`, dan `label_encoder.pkl` untuk memudahkan identifikasi dan penggunaan kembali di tahap implementasi.



```
[24] # 11. Simpan Model & Encoder
      joblib.dump(best_rf, "model_rf.pkl")
      joblib.dump(best_xgb, "model_xgb.pkl")
      joblib.dump(le, "label_encoder.pkl")

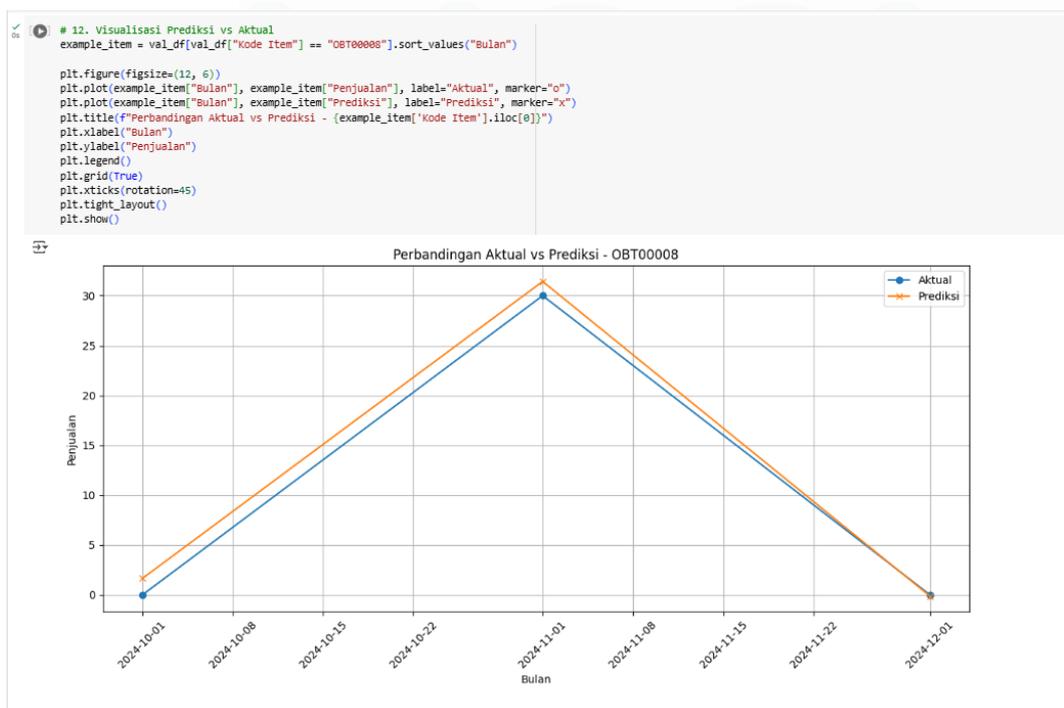
      ['label_encoder.pkl']
```

Gambar 3.16 Model dan Encoder

Dengan menyimpan model dan encoder secara permanen dalam format `.pkl`, pengguna dapat dengan mudah melakukan proses prediksi tanpa perlu memuat ulang dataset dan melakukan pelatihan dari awal. Hal ini sangat bermanfaat dalam skenario operasional di dunia nyata, misalnya dalam sistem prediksi stok obat yang

membutuhkan respon cepat dan efisien. Selain itu, file `label_encoder.pkl` juga penting karena model hanya akan mengenali data input dalam format numerik yang sama seperti saat dilatih, sehingga encoder ini wajib dimuat ulang ketika melakukan prediksi di kemudian hari. Dengan langkah ini, seluruh pipeline *machine learning* menjadi lebih modular, fleksibel, dan siap untuk diintegrasikan ke dalam sistem.

Gambar 3.18 menampilkan proses visualisasi perbandingan antara nilai aktual dan nilai prediksi penjualan obat untuk item dengan kode “OBT00008”. Visualisasi ini dilakukan dengan menggunakan pustaka `matplotlib` dalam Python. Data difilter berdasarkan Kode Item tertentu dan kemudian diurutkan berdasarkan kolom waktu “Bulan” agar garis tren terlihat secara kronologis. Ini merupakan langkah penting dalam mengevaluasi kualitas prediksi model dengan cara membandingkan prediksi terhadap data aktual yang telah terjadi.



Gambar 3.17 Perbandingan Data Aktual dan Prediksi

Plot yang dihasilkan terdiri dari dua garis, masing-masing mewakili data aktual dan prediksi. Garis biru melambangkan penjualan aktual dan diberi penanda lingkaran (“o”), sedangkan garis oranye menunjukkan hasil prediksi dengan

penanda silang (“x”). Kedua garis tersebut diplot terhadap sumbu x (bulan) dan sumbu y (jumlah penjualan), sehingga dapat dilihat bagaimana nilai prediksi mengikuti tren data aktual. Adanya kemiripan pola antara kedua garis menunjukkan bahwa model prediktif mampu menangkap tren musiman atau fluktuasi permintaan dengan baik.

Judul grafik secara otomatis disesuaikan berdasarkan Kode Item yang sedang divisualisasikan, dalam hal ini “OBT00008”. Hal ini membantu pengguna untuk langsung mengetahui konteks data yang sedang ditampilkan tanpa perlu melihat tabel atau metadata tambahan. Label sumbu x dan y juga disesuaikan agar lebih deskriptif, yaitu “Bulan” dan “Penjualan”, untuk memudahkan interpretasi bagi pengguna non-teknis. Penambahan elemen grid, rotasi label sumbu x, dan pengaturan layout yang rapi menunjukkan bahwa visualisasi ini dirancang agar nyaman dilihat dan mudah dipahami. Deskripsi lebih lanjut mengenai komponen dalam visualisasi Gambar 3.18 dapat dilihat pada Tabel 3.6.

Tabel 3.6 Deskripsi Komponen Visualisasi Perbandingan

Komponen	Deskripsi
Jenis Grafik	Garis dua seri
Garis Biru	Mewakili penjualan aktual, ditampilkan dengan penanda lingkaran (o)
Garis Oranye	Mewakili hasil prediksi model, ditampilkan dengan penanda silang (x)
Sumbu X	Bulan (periode waktu), menunjukkan dimensi waktu prediksi
Sumbu Y	Jumlah penjualan
Judul Grafik	Disesuaikan berdasarkan Kode Item yang sedang divisualisasikan, misalnya: “Prediksi Penjualan Item OBT00008”
Fungsi Utama	Membandingkan pola penjualan aktual dan prediksi serta menilai keakuratan model dalam menangkap tren dan fluktuasi permintaan.
Kegunaan bagi <i>Stakeholder</i>	Validasi prediksi stok secara visual dan Komunikasi hasil model secara intuitif kepada pihak non-teknis seperti manajer farmasi/pengadaan obat.

Visualisasi semacam ini sangat berguna dalam proses validasi model, terutama untuk stakeholder seperti manajer farmasi atau tim pengadaan obat. Dengan melihat seberapa dekat prediksi mengikuti data aktual, pengguna dapat

menilai akurasi dan keandalan model dalam perencanaan stok. Jika terdapat penyimpangan signifikan, grafik ini akan membantu dalam mengidentifikasi bulan atau pola musiman yang tidak berhasil ditangkap oleh model. Oleh karena itu, visualisasi ini tidak hanya menjadi alat analitik, tetapi juga sarana komunikasi hasil prediksi yang efektif kepada pihak non-teknis.

Tabel 3.7 memperlihatkan tampilan yang merupakan hasil akhir dari proses prediksi kebutuhan stok obat selama satu tahun. Setiap baris mewakili satu item obat yang diidentifikasi melalui kolom "Kode Item", sedangkan kolom-kolom selanjutnya menunjukkan hasil prediksi stok untuk masing-masing bulan dari Januari hingga Feb 2025. Untuk setiap bulan, tersedia tiga jenis nilai: nilai prediksi utama (Prediksi), nilai batas bawah (Prediksi_Min), dan nilai batas atas (Prediksi_Max). Ketiganya berguna untuk memberikan estimasi dengan margin kesalahan yang wajar berdasarkan performa model prediktif.

Tabel 3.7 Hasil Prediksi Algoritma XGBoost

Kode Item	Prediksi Jan 2025	Prediksi Min Jan 2025	Prediksi Max Jan 2025	Prediksi Feb 2025	Prediksi Min Feb 2025	Prediksi Max Feb 2025
OBT00001	1	1	27	2	1	28
OBT00004	135	109	161	130	104	156
OBT00006	32	6	58	54	28	80
OBT00008	1	1	27	7	1	33
OBT00016	39	13	65	31	5	57

Struktur data dalam Excel ini memudahkan pihak manajemen logistik rumah sakit dalam membaca dan menganalisis kebutuhan stok secara komprehensif. Dengan format ini, pengguna dapat membandingkan kebutuhan stok antarbulan, serta memahami fluktuasi dan tren permintaan yang diperkirakan. Selain itu, nilai minimum dan maksimum memungkinkan pengambilan keputusan yang lebih adaptif, misalnya menyesuaikan pembelian berdasarkan skenario

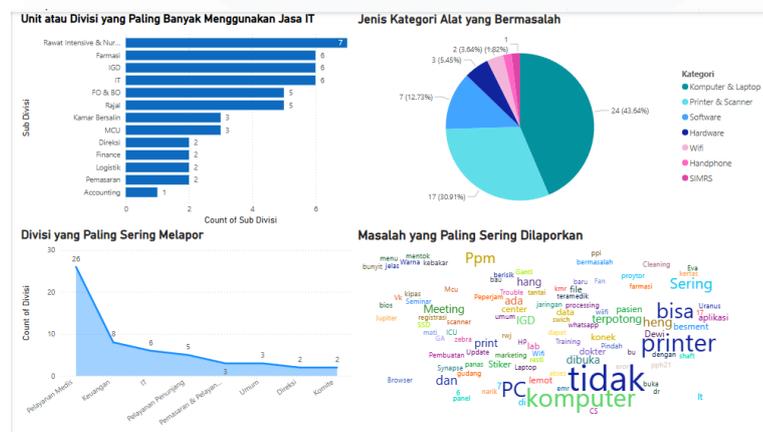
pesimis atau optimis. Format tabel seperti ini juga memudahkan integrasi ke dalam sistem manajemen stok.

File ini dapat digunakan sebagai alat bantu dalam proses pengadaan, pengelolaan gudang, dan perencanaan distribusi obat di Rumah Sakit Mentari. Informasi yang disediakan akan membantu mencegah kekurangan maupun kelebihan stok, yang dapat berakibat pada pemborosan atau ketidaksiapan dalam pelayanan. Selain itu, data ini juga relevan bagi divisi keuangan untuk menganggarkan biaya pembelian obat berdasarkan kebutuhan yang diprediksi secara kuantitatif. Dengan dukungan model prediktif dan penyajian data yang terstruktur seperti ini, rumah sakit dapat meningkatkan efisiensi dan ketepatan dalam pengelolaan obat sepanjang tahun 2025.

3.2.5 Mengembangkan Visualisasi Data dan *Dashboard* Interaktif

Dalam proses analisis data, pengembangan visualisasi dan *dashboard* interaktif merupakan langkah penting untuk menyajikan informasi secara lebih jelas dan terstruktur. Visualisasi ini membantu menjembatani hasil analisis teknis agar lebih mudah dipahami oleh pemangku kepentingan non-teknis. Tahap ini juga merupakan bagian dari proses *Modeling* dalam pendekatan CRISP-DM, yaitu *dashboard* dirancang menggunakan Power BI untuk menggabungkan berbagai metrik kinerja dan indikator utama dalam satu tampilan yang komprehensif. Visualisasi data ini mencakup grafik, tabel, serta elemen interaktif yang memungkinkan pengguna untuk mengeksplorasi tren layanan dan indikator kinerja secara lebih mendalam. Dengan adanya *dashboard* ini, pemangku kepentingan dapat dengan cepat mengidentifikasi pola, membandingkan data historis, serta memahami insight yang dihasilkan dari analisis data. Berikut merupakan beberapa visualisasi dan *dashboard* yang bertujuan untuk memperjelas temuan analisis serta memberikan gambaran menyeluruh terkait performa layanan di berbagai divisi.

Pada tahap ini, pengembangan visualisasi data dan *dashboard* interaktif menjadi elemen penting dalam proses analisis karena mampu menyajikan informasi secara lebih jelas, terstruktur, dan mudah dipahami oleh pemangku kepentingan. Visualisasi ini dirancang menggunakan Power BI, yang memungkinkan integrasi berbagai metrik dan indikator kinerja ke dalam satu tampilan *dashboard* yang komprehensif. Sebagaimana ditampilkan pada Gambar 3.19, *dashboard* ini dibuat berdasarkan dataset laporan layanan IT yang mencerminkan aktivitas dan kebutuhan teknis dari berbagai divisi dalam organisasi.



Gambar 3. 18 Visualisasi Laporan IT

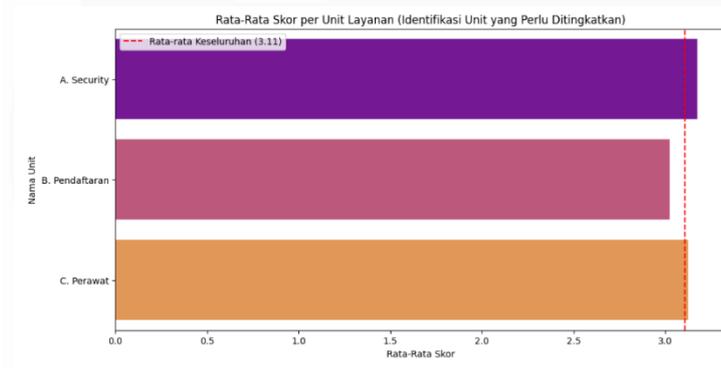
Visualisasi bar chart menunjukkan bahwa Sub Divisi Rawat Intensif merupakan unit yang paling banyak menggunakan layanan IT, diikuti oleh Farmasi, IGD, dan Kamar Bersalin. Hal ini mengindikasikan bahwa dukungan teknologi di unit-unit tersebut sangat krusial, terutama untuk operasional harian yang sangat bergantung pada perangkat keras dan sistem informasi. Informasi ini dapat digunakan oleh manajemen untuk menentukan prioritas dalam distribusi sumber daya teknis atau penguatan dukungan IT di area-area yang lebih sibuk.

Sementara itu, diagram pie memberikan gambaran proporsi masalah berdasarkan jenis perangkat, dengan hasil menunjukkan bahwa perangkat Komputer & Laptop menjadi penyumbang laporan terbanyak, disusul oleh Printer & Scanner serta perangkat lunak (Software). Grafik garis turut

memperlihatkan kecenderungan laporan dari waktu ke waktu, dengan Divisi Pelayanan Medis menjadi pelapor terbanyak, yang menunjukkan tingginya ketergantungan divisi ini terhadap sistem IT untuk operasional pelayanan. Wordcloud yang ditampilkan pada *dashboard* juga menjadi pelengkap yang membantu mengidentifikasi kata-kata yang sering muncul, seperti “tidak”, “printer”, “komputer”, dan “bisa”, yang memperkuat bukti bahwa permasalahan yang sering terjadi berkaitan dengan gangguan perangkat keras. Kombinasi berbagai jenis visualisasi ini memberikan pandangan holistik terhadap kondisi layanan IT dari berbagai dimensi.

Dashboard ini tidak hanya bersifat statis, tetapi juga dirancang secara interaktif agar pengguna dapat mengeksplorasi data lebih dalam berdasarkan kriteria tertentu, seperti divisi, jenis perangkat, atau periode waktu. Fitur interaktivitas ini memungkinkan pemangku kepentingan untuk menyesuaikan tampilan sesuai dengan kebutuhan analisis mereka, sehingga membantu dalam mengidentifikasi tren, permasalahan, maupun peluang perbaikan secara lebih cepat. Dalam konteks manajemen layanan IT, kemampuan untuk melihat dan menganalisis data secara real-time juga membantu pengambilan keputusan yang lebih responsif dan berbasis bukti. Oleh karena itu, *dashboard* ini tidak hanya berfungsi sebagai alat pelaporan, tetapi juga sebagai instrumen strategis dalam perencanaan dan peningkatan kualitas layanan secara berkelanjutan.

Selanjutnya, pada Gambar 3.20 ditampilkan visualisasi dari hasil pengolahan data kuesioner survei yang menunjukkan rata-rata skor layanan dari masing-masing unit. Visualisasi ini dirancang untuk membantu proses identifikasi terhadap unit-unit yang memiliki tingkat kepuasan pengguna di bawah rata-rata keseluruhan. Dalam grafik batang horizontal tersebut, ditampilkan tiga unit layanan, yaitu Unit Security, Unit Pendaftaran, dan Unit Perawat. Garis putus-putus merah vertikal menunjukkan rata-rata keseluruhan yang diperoleh, yaitu sebesar 3.11. Visualisasi ini juga menggunakan pewarnaan berbeda untuk setiap unit, guna memberikan pemisahan visual yang jelas antar unit.



Gambar 3.19 Visualisasi Identifikasi Unit

Setiap batang pada grafik diberi warna yang berbeda, yang tidak hanya berfungsi sebagai pemisah visual antar unit, tetapi juga memperkuat kejelasan pesan data yang ingin disampaikan kepada pengguna. Warna yang digunakan dalam visualisasi ini berperan penting dalam mempermudah pembaca untuk mengidentifikasi unit-unit secara cepat dan efisien, tanpa perlu membaca label satu per satu. Dengan adanya pembeda warna yang jelas, pembaca dapat segera mengetahui unit mana yang menunjukkan performa lebih baik atau lebih buruk dibandingkan dengan yang lain. Hal ini sangat berguna terutama dalam konteks presentasi yang memerlukan analisis cepat atau pemantauan oleh manajemen.

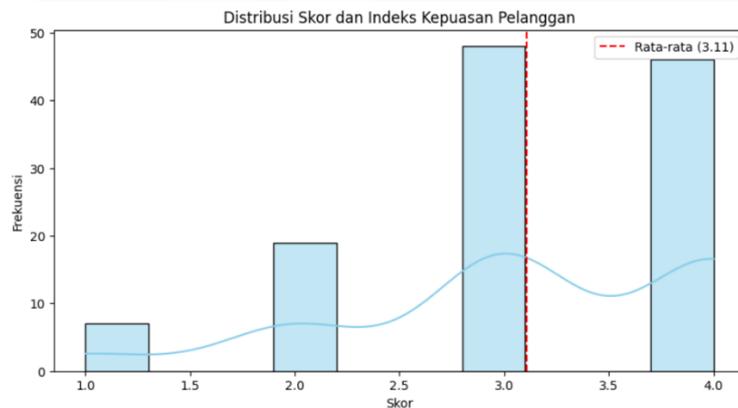
Dari hasil yang ditampilkan, terlihat bahwa Unit Pendaftaran dan Unit Perawat memiliki skor rata-rata yang berada di bawah nilai rata-rata keseluruhan, yang berarti tingkat kepuasan pengguna pada kedua unit ini masih belum optimal. Temuan ini menunjukkan adanya kebutuhan mendesak untuk melakukan perbaikan, baik dari aspek teknis maupun kualitas pelayanan. Perbaikan tersebut bisa melibatkan pelatihan bagi staf, evaluasi terhadap prosedur layanan yang diterapkan, ataupun penerapan teknologi baru untuk meningkatkan efisiensi dan responsivitas layanan. Menyadari adanya unit yang membutuhkan perhatian lebih ini, organisasi dapat merencanakan langkah-langkah yang lebih terfokus untuk mengatasi permasalahan yang ada.

Sementara itu, Unit Security memperoleh skor yang sedikit lebih tinggi dari rata-rata, yang menunjukkan bahwa kinerja unit ini sudah cukup baik namun masih ada ruang untuk perbaikan. Skor yang lebih tinggi ini bisa dijadikan acuan bagi unit-unit lain untuk mengadopsi beberapa praktik terbaik yang diterapkan oleh Unit Security. Oleh karena itu, temuan ini juga memberikan peluang untuk menggali lebih dalam mengenai faktor-faktor yang mendukung keberhasilan Unit Security, seperti pelatihan karyawan atau penggunaan teknologi untuk meningkatkan keamanan. Dengan mengetahui apa yang telah berhasil di unit tertentu, dapat diterapkan strategi serupa di unit lainnya untuk meningkatkan kinerja secara keseluruhan.

Visualisasi ini menjadi bagian penting dalam proses analisis karena mampu menyederhanakan informasi hasil survei dalam bentuk yang mudah dipahami oleh berbagai pihak, termasuk manajemen non-teknis. Dengan representasi visual yang jelas, pemangku kepentingan dapat secara cepat mengenali unit mana yang membutuhkan perhatian lebih besar dalam perbaikan layanan. Hal ini sejalan dengan tahapan *Modeling* dalam pendekatan CRISP-DM, di mana hasil analisis data disajikan dalam bentuk yang komunikatif dan actionable untuk mendukung pengambilan keputusan strategis. Oleh karena itu, visualisasi seperti ini tidak hanya berperan sebagai alat presentasi data, melainkan sebagai sarana utama dalam mengarahkan langkah-langkah peningkatan mutu yang berbasis pada bukti konkret.

Kemudian, Gambar 3.21 memperlihatkan distribusi skor kepuasan pelanggan berdasarkan hasil survei, dengan rentang skor antara 1 hingga 4. Dalam survei ini, skor 1 menunjukkan tingkat kepuasan yang sangat rendah atau "kurang sekali", skor 2 mengindikasikan bahwa layanan tersebut dianggap "kurang", skor 3 bermakna bahwa layanan tersebut "baik", dan skor 4 menunjukkan tingkat kepuasan yang sangat tinggi atau "baik sekali". Histogram ini memberikan gambaran visual yang jelas mengenai persepsi pelanggan terhadap layanan yang diberikan oleh organisasi. Grafik ini menjadi alat yang efektif dalam menganalisis tingkat kepuasan secara

keseluruhan, dan mempermudah untuk mengidentifikasi apakah ada kecenderungan kepuasan yang dominan atau area yang memerlukan perhatian khusus.



Gambar 3.20 Visualisasi Indeks Kepuasan Pelanggan

Dalam visualisasi ini, terlihat bahwa sebagian besar responden memberikan skor pada level 3 dan 4, yang berarti mayoritas pelanggan merasa bahwa layanan yang diberikan sudah berada dalam kategori baik hingga sangat baik. Hal ini menunjukkan bahwa layanan yang diberikan oleh organisasi pada umumnya telah memenuhi harapan pelanggan dan sudah dapat dianggap memadai atau lebih. Dominasi skor 3 dan 4 ini mengindikasikan bahwa organisasi telah berhasil dalam memberikan layanan yang memuaskan untuk sebagian besar pelanggan, yang tentunya merupakan pencapaian positif dalam hal kualitas layanan.

Namun, meskipun secara keseluruhan persepsi pelanggan cenderung positif, grafik ini juga memperlihatkan adanya responden yang memberikan skor 1 dan 2. Skor 1, yang termasuk dalam kategori "kurang sekali", dan skor 2, yang berarti "kurang", menunjukkan bahwa ada sebagian kecil pelanggan yang merasa sangat tidak puas atau belum puas dengan layanan yang diterima. Temuan ini penting untuk diidentifikasi, karena meskipun mayoritas pelanggan memberikan umpan balik positif, kehadiran skor rendah menandakan adanya sejumlah kecil pelanggan yang

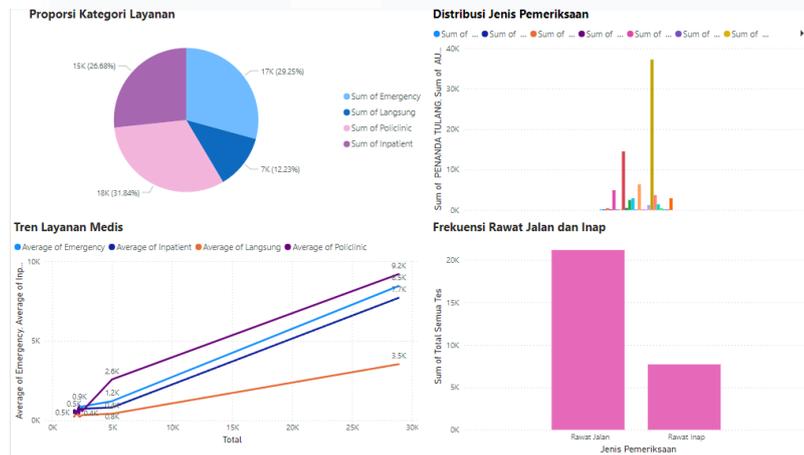
memiliki pengalaman buruk atau merasa ada aspek-aspek tertentu yang belum memenuhi harapan mereka.

Temuan tersebut menjadi dasar evaluasi yang penting, karena meskipun rata-rata kepuasan berada pada tingkat baik, masih ada ruang untuk perbaikan. Sebagai contoh, organisasi perlu menganalisis lebih lanjut faktor-faktor yang menyebabkan sebagian kecil pelanggan merasa tidak puas, dan langkah-langkah perbaikan dapat diterapkan untuk mengurangi pengalaman negatif tersebut. Dengan memahami penyebab ketidakpuasan, organisasi dapat menargetkan area-area yang perlu diperbaiki untuk meningkatkan pengalaman keseluruhan pelanggan. Dengan cara ini, organisasi dapat mengurangi skor rendah yang muncul dan meningkatkan tingkat kepuasan secara lebih menyeluruh.

Dengan demikian, visualisasi ini tidak hanya berfungsi sebagai alat untuk menggambarkan tingkat kepuasan secara keseluruhan, tetapi juga sebagai titik awal untuk perbaikan dan peningkatan lebih lanjut dalam layanan yang diberikan. Garis merah putus-putus vertikal yang menunjukkan rata-rata keseluruhan, yaitu 3,11, memberikan acuan yang jelas bagi organisasi untuk menentukan area yang perlu difokuskan dalam strategi peningkatan kualitas layanan. Temuan ini mendukung pendekatan berbasis data dalam proses pengambilan keputusan, sehingga organisasi dapat merencanakan perbaikan yang lebih terukur dan efektif untuk memenuhi ekspektasi pelanggan secara keseluruhan.

Gambar 3.22 menyajikan visualisasi data penggunaan layanan laboratorium yang dibagi dalam beberapa kategori analisis, yaitu proporsi kategori layanan, distribusi jenis pemeriksaan, tren layanan medis, dan frekuensi rawat jalan dan rawat inap. Diagram lingkaran di kiri atas memperlihatkan bahwa layanan laboratorium paling banyak digunakan oleh pasien poliklinik (31,84%) dan pasien yang datang dengan kondisi emergency (29,23%). Sementara itu, kategori inpatient (rawat inap) dan langsung (direct) memiliki proporsi yang lebih rendah. Temuan ini

menunjukkan bahwa mayoritas layanan laboratorium berfokus pada pasien rawat jalan dan kondisi gawat darurat, yang sering membutuhkan diagnosis cepat dan penanganan segera.



Gambar 3.21 Visualisasi Proporsi Layanan Laboratorium

Pada grafik batang yang terletak di kanan atas, distribusi jenis pemeriksaan laboratorium menunjukkan dominasi beberapa jenis tes, seperti tes hematologi dan kimia darah, yang paling banyak digunakan di kategori layanan emergency dan poliklinik. Hal ini mengindikasikan bahwa kedua kategori ini lebih banyak membutuhkan pemeriksaan cepat dan rutin untuk mendiagnosis kondisi pasien yang datang dengan keluhan medis yang mendesak atau yang membutuhkan kontrol kesehatan secara berkala. Dominasi tes-tes ini memberikan gambaran yang jelas mengenai jenis layanan laboratorium yang paling sering digunakan sesuai dengan jenis layanan medis yang diberikan.

Sementara itu, grafik garis di kiri bawah menggambarkan tren penggunaan layanan laboratorium berdasarkan kategori layanan. Grafik ini menunjukkan adanya peningkatan yang konsisten pada kategori emergency dan inpatient, seiring dengan meningkatnya jumlah layanan yang diberikan. Ini mencerminkan kecenderungan bahwa semakin banyak pasien yang membutuhkan pemeriksaan laboratorium di kondisi darurat atau rawat inap, yang menunjukkan bahwa penggunaan layanan laboratorium semakin

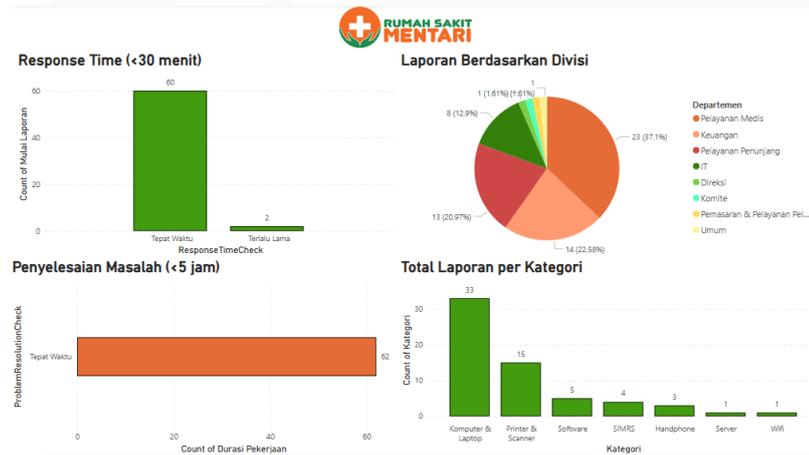
meningkat seiring dengan kebutuhan medis yang lebih mendesak. Di sisi lain, kategori langsung menunjukkan tren pertumbuhan yang lebih lambat, yang bisa jadi disebabkan oleh jenis pemeriksaan yang lebih terbatas dan tidak memerlukan proses diagnostik mendalam.

Pada grafik batang di kanan bawah, terlihat frekuensi pemeriksaan laboratorium yang jauh lebih tinggi pada layanan rawat jalan dibandingkan rawat inap. Hal ini menandakan bahwa pemeriksaan laboratorium lebih sering dilakukan untuk pasien yang menjalani perawatan rawat jalan, yang biasanya datang untuk pemeriksaan rutin atau tindak lanjut medis. Pasien rawat inap mungkin memiliki kebutuhan laboratorium yang lebih spesifik dan terbatas pada kondisi medis tertentu, sehingga frekuensinya lebih rendah. Dengan demikian, temuan ini memberi wawasan mengenai kebutuhan layanan laboratorium yang lebih tinggi pada pasien non-inap, yang lebih sering memerlukan pemeriksaan diagnostik secara rutin.

Secara keseluruhan, Gambar 3.22 memberikan gambaran menyeluruh tentang pemanfaatan layanan laboratorium yang lebih banyak terjadi pada layanan rawat jalan dan kondisi gawat darurat. Hal ini menunjukkan bahwa dukungan laboratorium sangat penting dalam mendukung diagnosa cepat, penanganan segera pasien di kondisi gawat darurat, dan pemeriksaan rutin untuk pasien non-inap. Temuan ini mengindikasikan bahwa fasilitas laboratorium perlu diprioritaskan dalam mendukung kedua kategori tersebut untuk memastikan kelancaran proses diagnostik dan meningkatkan kualitas pelayanan medis yang diberikan kepada pasien.

Gambar 3.23 menyajikan laporan capaian mutu layanan Teknologi Informasi (TI) di Rumah Sakit Mentari, yang dianalisis berdasarkan beberapa variabel kunci: waktu respons, waktu penyelesaian masalah, divisi pelapor, dan kategori masalah. Grafik-grafik yang ditampilkan dalam visualisasi ini memberikan gambaran yang jelas tentang performa layanan TI serta kualitas respons yang diberikan tim TI dalam menangani berbagai masalah teknis. Tujuan dari visualisasi ini adalah untuk mengidentifikasi

kinerja tim TI dalam memberikan layanan tepat waktu dan efisien, sekaligus mengetahui area yang perlu mendapatkan perhatian lebih untuk perbaikan ke depannya.



Gambar 3.22 Visualisasi Capaian Mutu Divisi IT

Pada grafik kiri atas, terlihat bahwa mayoritas laporan (60 laporan) berhasil ditangani dalam waktu respons kurang dari 30 menit, yang menunjukkan bahwa tim TI sangat responsif terhadap permasalahan yang terjadi. Hanya terdapat 2 laporan yang mengalami keterlambatan, yang masih berada dalam batas toleransi yang dapat diterima. Keberhasilan tim TI dalam menjaga waktu respons yang cepat ini menunjukkan tingkat profesionalisme yang tinggi dalam menangani permintaan layanan. Respons yang cepat sangat penting dalam mendukung kelancaran operasional rumah sakit, mengingat banyaknya perangkat teknis yang mendukung proses pelayanan medis yang harus berfungsi dengan optimal.

Selanjutnya, grafik kiri bawah menunjukkan waktu penyelesaian masalah, yang merupakan indikator penting lainnya untuk mengukur efisiensi tim TI. Sebagian besar masalah (62 laporan) diselesaikan dalam waktu kurang dari 5 jam, mencerminkan efisiensi yang baik dalam menangani dan menyelesaikan permasalahan teknis. Proses penyelesaian masalah yang cepat dan efisien sangat berkontribusi terhadap kelancaran operasional rumah sakit, mengingat banyaknya perangkat dan sistem yang

harus selalu siap digunakan tanpa gangguan. Tim TI yang mampu menyelesaikan masalah dengan cepat tidak hanya mendukung produktivitas rumah sakit tetapi juga meningkatkan kepuasan pengguna layanan.

Pada grafik pie di kanan atas, distribusi laporan berdasarkan divisi menunjukkan bahwa sebagian besar laporan berasal dari Divisi Pelayanan Penunjang (42,59%), diikuti oleh Divisi Keuangan dan Divisi Pelayanan Medis. Ini memberikan informasi penting bahwa layanan TI yang diberikan cenderung lebih banyak digunakan oleh unit-unit yang mendukung operasional harian rumah sakit, seperti layanan penunjang dan administrasi. Meskipun kategori pelayanan medis juga memberikan kontribusi yang signifikan terhadap jumlah laporan, data ini memberikan indikasi bahwa perangkat TI yang digunakan oleh unit penunjang sangat vital untuk menjaga kelancaran proses rumah sakit, seperti sistem manajemen rumah sakit, perangkat administrasi pasien, dan sistem pembukuan keuangan.

Selain itu, distribusi laporan berdasarkan kategori masalah yang terlihat pada grafik kanan bawah memberikan wawasan yang lebih mendalam tentang masalah teknis yang paling sering dihadapi oleh rumah sakit. Kategori masalah yang paling banyak dilaporkan adalah terkait dengan perangkat komputer dan laptop, dengan total 33 laporan. Ini menunjukkan bahwa perangkat keras masih menjadi area yang paling banyak mengalami kendala teknis, yang dapat menghambat proses operasional. Setelah komputer dan laptop, masalah terkait printer & scanner serta software menyusul dalam urutan frekuensi laporan yang tinggi. Hal ini menandakan bahwa meskipun perangkat keras masih menjadi masalah utama, gangguan terkait perangkat lunak dan periferal juga memberikan dampak yang signifikan terhadap kinerja layanan TI.

Data ini memberikan gambaran yang sangat berharga bagi tim TI dan manajemen rumah sakit dalam merencanakan langkah-langkah perbaikan dan pengembangan lebih lanjut. Berdasarkan temuan ini, tim TI perlu fokus pada peningkatan kualitas perangkat keras, termasuk pengadaan

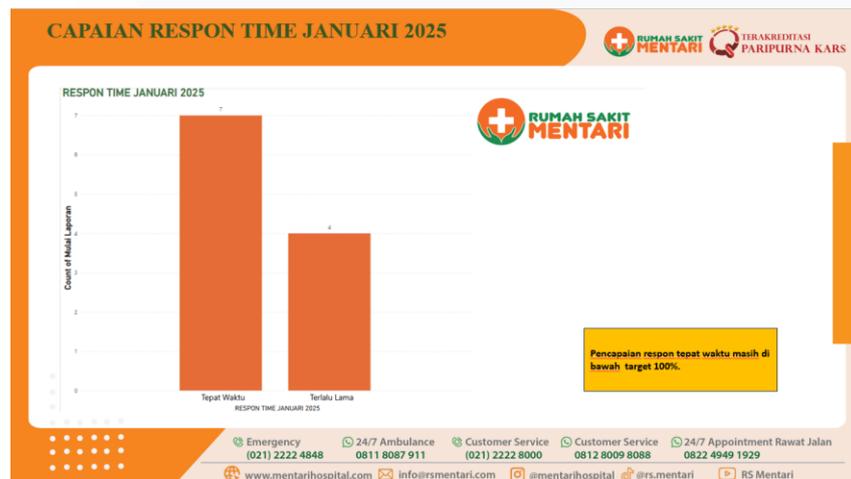
dan pemeliharaan perangkat komputer dan laptop yang lebih baik, serta peningkatan kapasitas teknis untuk menangani masalah perangkat lunak dan printer. Selain itu, temuan ini juga dapat menjadi dasar dalam merencanakan investasi di masa depan, dengan mempertimbangkan pengadaan perangkat yang lebih tahan lama dan efektif serta peningkatan kemampuan tim dalam menangani berbagai masalah teknis yang sering muncul.

Dengan melihat data yang ada, tim TI di Rumah Sakit Mentari dapat lebih fokus pada area yang membutuhkan perhatian lebih, seperti pemeliharaan dan penggantian perangkat keras yang sudah tidak memadai. Mengingat bahwa kebanyakan masalah berkaitan dengan perangkat keras, salah satu langkah yang dapat diambil adalah meningkatkan pengelolaan inventaris TI dan memastikan bahwa perangkat yang digunakan selalu dalam kondisi optimal. Selain itu, peningkatan pelatihan dan keterampilan teknis tim TI juga dapat membantu mengurangi waktu respons dan waktu penyelesaian masalah, memastikan bahwa rumah sakit dapat terus beroperasi dengan lancar meskipun ada gangguan teknis.

Secara keseluruhan, Gambar 3.23 memberikan gambaran komprehensif tentang capaian mutu layanan TI di Rumah Sakit Mentari. Dengan adanya visualisasi ini, manajemen rumah sakit dapat lebih mudah memantau kinerja tim TI dan mengidentifikasi area yang perlu perbaikan. Dengan mengimplementasikan rekomendasi berbasis data ini, rumah sakit dapat meningkatkan efisiensi operasional dan kualitas layanan yang diberikan kepada pasien serta seluruh staf rumah sakit. Data ini juga menjadi acuan bagi perencanaan jangka panjang dalam pengembangan infrastruktur TI yang lebih baik, untuk mendukung kelancaran operasional rumah sakit yang semakin kompleks.

Gambar 3.24 menyajikan visualisasi capaian waktu respons terhadap laporan yang ditangani oleh Divisi *General Affair* Rumah Sakit Mentari pada bulan Januari 2025. Grafik batang menunjukkan bahwa sebanyak 7 laporan berhasil ditangani tepat waktu, sementara 4 laporan

mengalami keterlambatan dalam penanganannya. Meskipun sebagian besar laporan ditindaklanjuti sesuai dengan target waktu yang ditentukan, pencapaian ketepatan waktu masih berada di bawah standar 100% yang diharapkan. Hal ini mengindikasikan bahwa meskipun respons yang diberikan sudah cukup baik, masih ada ruang untuk perbaikan dalam hal kecepatan dan ketepatan waktu dalam merespons laporan yang masuk.



Gambar 3.23 Visualisasi Capaian Respon Time General Affair

Pencapaian ini, meskipun positif, tetap menunjukkan adanya tantangan yang perlu diperbaiki oleh Divisi *General Affair*. Meskipun 7 laporan berhasil ditangani tepat waktu, angka 4 laporan yang terlambat memberikan respons bahwa ada area yang membutuhkan perhatian lebih. Keterlambatan dalam menangani laporan dapat mempengaruhi kelancaran operasional rumah sakit, terutama dalam hal pelayanan pendukung non-medis yang berperan penting untuk mendukung berbagai aktivitas medis dan administrasi. Oleh karena itu, penting bagi Divisi *General Affair* untuk mengidentifikasi faktor-faktor yang menyebabkan keterlambatan ini.

Visualisasi ini menekankan bahwa pencapaian capaian waktu respons yang belum sepenuhnya optimal perlu dievaluasi lebih lanjut. Untuk mendorong perbaikan mutu pelayanan, Divisi *General Affair* dapat melakukan analisis mendalam terhadap alur penanganan laporan yang ada.

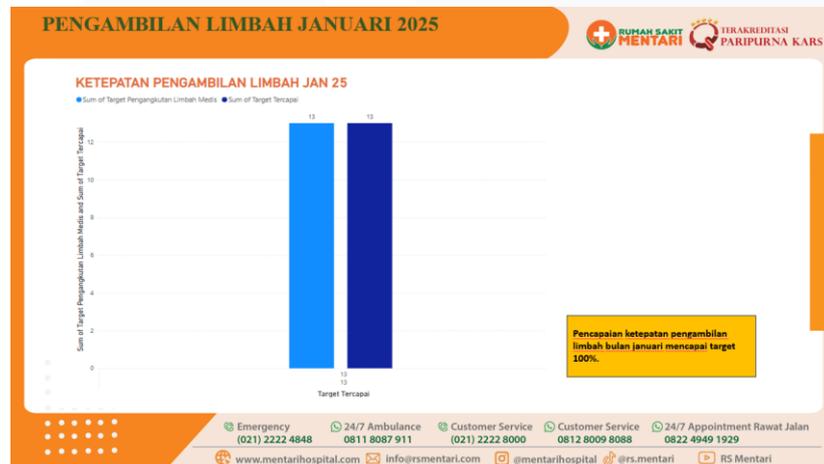
Dengan memetakan proses yang terlibat dalam penanganan laporan, tim dapat mengidentifikasi titik-titik keterlambatan dan potensi masalah dalam sistem. Hal ini akan memberikan gambaran yang jelas tentang langkah-langkah yang dapat diambil untuk mempercepat proses penanganan laporan di masa mendatang.

Selain itu, peningkatan kapasitas sumber daya manusia juga menjadi faktor yang krusial. Pelatihan dan pengembangan keterampilan tim dalam menangani laporan atau permintaan yang masuk dapat membantu meningkatkan kecepatan dan akurasi respons. Sistem pelaporan yang lebih efisien dan transparan juga dapat menjadi solusi yang mendukung dalam meningkatkan waktu respons. Pembaruan atau peningkatan perangkat lunak untuk pelaporan atau pengelolaan data laporan juga bisa menjadi langkah strategis untuk meningkatkan kualitas pelayanan internal rumah sakit.

Secara keseluruhan, visualisasi capaian waktu respons yang ditampilkan dalam Gambar 3.24 memberikan gambaran yang jelas mengenai efektivitas tim Divisi *General Affair* dalam menangani laporan. Meskipun sebagian besar laporan ditangani tepat waktu, pencapaian yang sedikit di bawah standar ini menunjukkan bahwa perbaikan masih diperlukan. Dengan melaksanakan evaluasi menyeluruh terhadap alur penanganan laporan dan meningkatkan kapasitas sumber daya manusia serta sistem pelaporan, diharapkan Divisi *General Affair* dapat mencapai tingkat pelayanan yang lebih optimal dan mendukung kelancaran operasional rumah sakit secara keseluruhan.

Gambar 3.25 menampilkan visualisasi capaian mutu Divisi *General Affair* di Rumah Sakit Mentari untuk indikator ketepatan pengambilan limbah medis pada bulan Januari 2025. Dalam grafik batang yang ditampilkan, terlihat bahwa jumlah target pengambilan limbah sebanyak 13 kali tercapai sepenuhnya, dengan ketinggian batang yang sama antara target dan realisasi. Hal ini mencerminkan bahwa seluruh aktivitas pengambilan limbah telah dilakukan sesuai dengan jadwal yang telah ditetapkan tanpa

ada keterlambatan atau kegagalan. Dengan pencapaian ini, tim *General Affair* berhasil mencapai tingkat ketepatan pengambilan limbah sebesar 100%, yang menunjukkan kinerja yang sangat baik.



Gambar 3.24 Visualisasi Capaian Ketepatan Pengambilan Limbah Medis

Capaian 100% ini menjadi indikasi bahwa pengelolaan limbah medis di Rumah Sakit Mentari dilakukan dengan sangat disiplin dan sesuai dengan standar yang telah ditetapkan. Proses pengambilan limbah medis merupakan salah satu aspek yang sangat penting dalam menjaga kebersihan dan keamanan lingkungan rumah sakit. Dengan pengelolaan yang tepat waktu dan konsisten, risiko kontaminasi atau masalah kesehatan terkait limbah medis dapat diminimalkan. Keberhasilan ini menunjukkan komitmen divisi terkait terhadap aspek keselamatan dan kebersihan rumah sakit, yang sangat vital dalam mendukung proses perawatan pasien.

Visualisasi ini tidak hanya memperlihatkan pencapaian yang baik dari sisi ketepatan waktu, tetapi juga menggambarkan efektivitas koordinasi internal dalam tim Divisi *General Affair*. Pencapaian ini menunjukkan bahwa seluruh anggota tim mampu bekerja sesuai dengan prosedur operasional standar (SOP) yang ada, serta melaksanakan tugas mereka dengan penuh tanggung jawab. Hal ini menunjukkan kepatuhan yang tinggi terhadap protokol yang telah ditetapkan, yang berperan penting dalam menjaga kualitas lingkungan rumah sakit.

Pencapaian 100% pada indikator ini juga memberikan gambaran mengenai pentingnya perencanaan yang matang dan alur komunikasi yang efisien antara tim *General Affair* dengan unit-unit lainnya di rumah sakit. Ketepatan dalam pengambilan limbah medis tidak hanya bergantung pada jadwal yang ditentukan, tetapi juga pada kesiapan dan koordinasi antar tim. Dengan memperhatikan detail-detail kecil seperti jadwal pengambilan, volume limbah, dan pemilahan limbah yang benar, proses ini dapat berjalan dengan lancar tanpa hambatan.

Keberhasilan Divisi *General Affair* dalam pengelolaan limbah medis juga memberikan landasan yang kuat untuk memperluas pemantauan mutu pada indikator lainnya. Misalnya, pengelolaan linen, distribusi aset, serta layanan fasilitas umum di rumah sakit. Pengalaman positif dalam mencapai target 100% untuk pengambilan limbah medis dapat menjadi contoh bagi indikator lain dalam upaya memperbaiki dan meningkatkan kualitas operasional rumah sakit secara keseluruhan. Hal ini menunjukkan bahwa Divisi *General Affair* dapat meningkatkan efisiensi dan efektivitas di berbagai bidang.

Dalam konteks manajemen mutu, pencapaian ini lebih dari sekadar angka. Ini mencerminkan keseriusan Rumah Sakit Mentari dalam menjaga kualitas operasional yang berkaitan langsung dengan keselamatan dan kenyamanan pasien. Mengelola limbah medis secara tepat waktu dan sesuai prosedur bukan hanya merupakan kewajiban administratif, tetapi juga bagian dari upaya rumah sakit untuk menjaga reputasi dan kredibilitasnya sebagai institusi kesehatan yang bertanggung jawab.

Dengan demikian, capaian 100% dalam ketepatan pengambilan limbah medis pada Januari 2025 ini tidak hanya menunjukkan keberhasilan teknis, tetapi juga merupakan indikator kinerja yang mendukung keseluruhan kualitas layanan di Rumah Sakit Mentari. Keberhasilan ini menunjukkan bahwa tim *General Affair* mampu menjalankan tugas dengan baik dan berperan besar dalam menjaga lingkungan yang aman dan bersih,

yang pada akhirnya mendukung keseluruhan pelayanan kesehatan kepada pasien. Pencapaian ini juga menjadi fondasi bagi perbaikan lebih lanjut dalam area lainnya, guna memastikan rumah sakit dapat terus memberikan pelayanan terbaik bagi pasien dan masyarakat.

Gambar 3.26 menampilkan visualisasi distribusi pekerjaan yang dilakukan oleh Divisi *General Affair* Rumah Sakit Mentari pada bulan Januari 2025 berdasarkan kategori tugas. Grafik lingkaran ini memberikan gambaran proporsional atas jenis pekerjaan yang paling banyak dilakukan sepanjang bulan tersebut. Berdasarkan visualisasi, terlihat bahwa pekerjaan kategori Elektrik mendominasi dengan persentase sebesar 63,64%, menjadikannya kategori tertinggi dalam daftar pekerjaan.



Gambar 3.25 Distribusi Pekerjaan General Affair

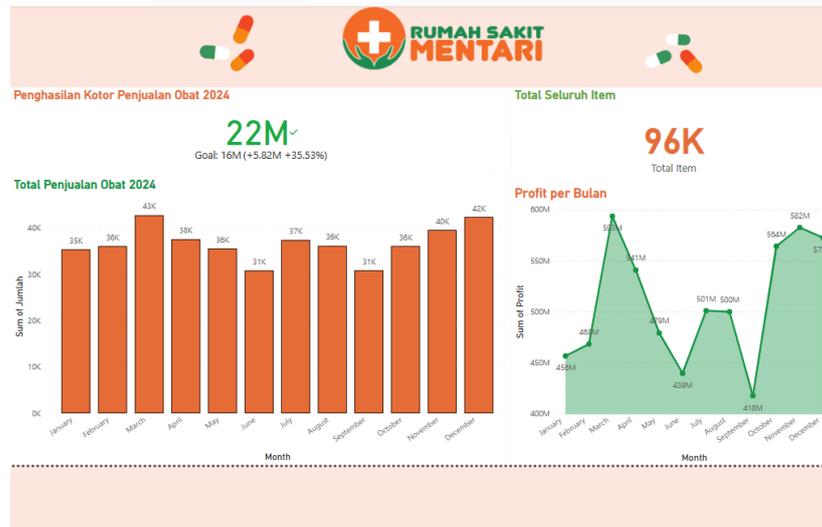
Kategori pekerjaan lain yang juga cukup signifikan adalah Plumbing, yang mencakup 21,21% dari total pekerjaan. Sementara itu, pekerjaan terkait Peralatan Medis, Sipil, dan Telepon masing-masing mencakup 6,06%, 6,06%, dan 3,03%. Komposisi ini menunjukkan bahwa sebagian besar aktivitas perbaikan dan pemeliharaan yang dilakukan oleh tim *General Affair* berfokus pada sistem kelistrikan dan saluran air, dua aspek penting dalam mendukung operasional rumah sakit yang stabil dan aman.

Distribusi ini dapat mencerminkan kebutuhan teknis dan infrastruktur rumah sakit selama periode tersebut. Tingginya proporsi pekerjaan elektrik bisa menandakan adanya intensitas penggunaan alat-alat listrik dan sistem penerangan yang memerlukan pemantauan serta penanganan cepat demi menjamin kelancaran layanan kesehatan. Demikian pula, pekerjaan plumbing yang cukup dominan mencerminkan perhatian terhadap aspek sanitasi dan kenyamanan lingkungan rumah sakit.

Dengan adanya klasifikasi pekerjaan seperti ini, manajemen dapat lebih mudah melakukan evaluasi dan perencanaan kerja ke depan. Misalnya, jika tren serupa berulang dalam beberapa bulan berturut-turut, maka akan dibutuhkan alokasi sumber daya yang lebih besar untuk kategori pekerjaan tersebut. Ini juga menjadi indikator penting dalam pengelolaan SDM dan kebutuhan pelatihan teknis bagi staf *General Affair* agar tetap responsif dan terampil dalam menangani jenis pekerjaan yang paling sering terjadi.

Secara keseluruhan, visualisasi distribusi pekerjaan ini tidak hanya memberikan informasi statistik, tetapi juga mendukung pengambilan keputusan strategis dalam pengelolaan fasilitas rumah sakit. Dengan memahami area mana yang paling membutuhkan perhatian teknis, rumah sakit dapat meningkatkan efisiensi operasional serta menjaga standar kualitas layanan yang tinggi bagi seluruh pasien dan pengguna fasilitas.

Selanjutnya, Gambar 3.27 menggambarkan performa penjualan obat oleh unit farmasi Rumah Sakit Mentari sepanjang tahun 2024. Total penjualan obat mencapai 22 juta rupiah, melampaui target tahunan yang telah ditetapkan sebesar 16 juta rupiah, dengan selisih positif 5,82 juta rupiah atau 35,53% di atas target. Pencapaian ini menunjukkan kinerja yang sangat baik dari unit farmasi dalam memenuhi permintaan obat-obatan di rumah sakit dan mencapai hasil yang lebih baik dari yang direncanakan. Selisih positif ini menjadi indikator bahwa manajemen persediaan obat dan strategi pemasaran farmasi berjalan dengan efektif.



Gambar 3.26 Visualisasi Perfoma Penjualan Obat Farmasi

Jumlah item yang terjual secara keseluruhan mencapai 96.000 unit, menunjukkan bahwa permintaan terhadap obat-obatan di rumah sakit relatif tinggi dan stabil sepanjang tahun. Grafik batang yang menampilkan penjualan bulanan menunjukkan bahwa penjualan tertinggi tercatat pada bulan Februari dan Desember, dengan masing-masing mencapai 43 ribu item. Pencapaian ini dapat disebabkan oleh faktor-faktor tertentu seperti peningkatan jumlah pasien atau permintaan musiman terkait penyakit tertentu yang memerlukan pengobatan intensif selama periode tersebut. Hal ini menunjukkan adanya permintaan yang cukup tinggi dan konsisten untuk obat-obatan di rumah sakit.

Namun, meskipun ada lonjakan penjualan pada beberapa bulan, bulan Juli mencatatkan penjualan terendah, hanya mencapai 31 ribu item. Penurunan penjualan ini dapat dipengaruhi oleh beberapa faktor, seperti periode liburan, penurunan jumlah kunjungan pasien, atau pengurangan jumlah obat yang dibutuhkan untuk pengobatan tertentu. Fluktuasi penjualan seperti ini dapat menjadi bahan pertimbangan dalam merencanakan strategi pengadaan dan distribusi obat ke depannya, guna menjaga agar pasokan obat tetap stabil dan mencakup seluruh kebutuhan pasien.

Stabilitas volume penjualan yang relatif terjaga dari bulan ke bulan juga mencerminkan adanya permintaan yang konsisten terhadap obat-obatan di rumah sakit. Walaupun terjadi fluktuasi di beberapa bulan, pola umum menunjukkan bahwa permintaan untuk obat tetap ada sepanjang tahun. Hal ini menunjukkan pentingnya keberlanjutan dalam penyediaan obat yang cukup, yang mendukung kelancaran proses perawatan pasien tanpa terganggu oleh kelangkaan atau kelebihan stok.

Selain itu, grafik garis yang menggambarkan laba menunjukkan fluktuasi yang cukup signifikan dari bulan ke bulan. Puncak laba terjadi pada bulan Maret, yang mencatatkan laba sekitar 580 juta rupiah, sementara bulan Desember juga mencatatkan laba yang cukup tinggi, yaitu 572 juta rupiah. Pencapaian laba tertinggi ini dapat berkaitan dengan penjualan obat yang meningkat pada periode-periode tertentu, baik karena volume penjualan yang lebih tinggi maupun harga obat yang lebih menguntungkan pada bulan-bulan tersebut.

Di sisi lain, penurunan laba paling tajam terjadi pada bulan September, dengan total profit hanya mencapai 418 juta rupiah. Penurunan laba yang cukup signifikan ini bisa disebabkan oleh beberapa faktor, seperti fluktuasi harga obat, perbedaan jenis obat yang terjual pada periode tersebut, atau strategi harga yang mungkin kurang optimal. Fluktuasi laba ini perlu dianalisis lebih dalam untuk memahami faktor-faktor penyebabnya, sehingga dapat dilakukan perbaikan pada strategi penjualan dan pengadaan obat.

Perbedaan laba antara bulan-bulan tertentu mencerminkan dinamika pengadaan obat dan strategi harga yang diterapkan oleh unit farmasi. Misalnya, jika bulan-bulan dengan laba tinggi didominasi oleh penjualan obat-obat dengan margin keuntungan lebih tinggi, sementara bulan dengan laba rendah didominasi oleh obat-obat dengan margin yang lebih rendah, maka ini dapat menjadi petunjuk untuk merencanakan strategi penjualan yang lebih menguntungkan di masa depan. Pengelolaan harga dan

pengadaan yang tepat dapat membantu mengoptimalkan laba tanpa mengurangi kualitas layanan kepada pasien.

Secara umum, performa penjualan farmasi Rumah Sakit Mentari pada tahun 2024 menunjukkan pencapaian yang positif dan melampaui target yang telah ditetapkan. Keberhasilan ini menjadi dasar yang kuat untuk merencanakan strategi pengadaan obat dan penjualan ke depan. Meskipun ada fluktuasi dalam laba dan volume penjualan di beberapa bulan, pencapaian ini mencerminkan adanya permintaan yang konsisten terhadap obat-obatan di rumah sakit dan kemampuan unit farmasi untuk memenuhi kebutuhan tersebut dengan efektif. Data ini dapat digunakan untuk perencanaan keuangan dan operasional di tahun-tahun mendatang, serta dalam menyusun strategi untuk mempertahankan atau meningkatkan performa penjualan obat. Oleh karena itu, penting bagi manajemen farmasi untuk terus memantau tren penjualan dan profitabilitas secara berkala guna mengidentifikasi peluang peningkatan serta mengantisipasi potensi risiko di masa depan.

3.2.6 Memberikan Rekomendasi Berbasis Data

Setelah proses analisis data selesai dilakukan, langkah selanjutnya adalah menyusun rekomendasi berbasis data. Rekomendasi ini dirumuskan dengan mengidentifikasi pola dan temuan penting yang memengaruhi kinerja layanan di Rumah Sakit Mentari. Misalnya, dari hasil visualisasi penjualan obat, dapat diberikan rekomendasi untuk menyesuaikan stok berdasarkan tren bulanan agar tidak terjadi kelebihan atau kekurangan pasokan. Di sisi lain, analisis laporan TI dan *General Affair* dapat menghasilkan usulan peningkatan kecepatan respons atau peninjauan ulang terhadap sistem pelaporan internal. Rekomendasi ini bersifat praktis dan kontekstual karena langsung didasarkan pada data yang merefleksikan kondisi aktual rumah sakit.

Dalam konteks metodologi CRISP-DM (*Cross-Industry Standard Process for Data mining*), bagian ini masuk dalam tahapan Evaluation. Pada fase ini, hasil analisis dievaluasi bersama dengan tim rumah sakit untuk memastikan bahwa temuan yang diperoleh benar-benar relevan dengan tujuan bisnis dan dapat ditindaklanjuti. Proses evaluasi ini juga menjadi sarana validasi terhadap kualitas data dan interpretasi analisis yang dilakukan. Hasilnya digunakan untuk menentukan langkah selanjutnya, termasuk apakah diperlukan analisis tambahan atau sudah cukup untuk langsung berlanjut ke tahap Deployment. Evaluasi bersama ini penting untuk memastikan bahwa rekomendasi yang diberikan tidak hanya akurat dari sisi data, tetapi juga realistis dan aplikatif dalam konteks operasional rumah sakit.

3.2.7 Menyarankan Inisiatif Perbaikan Berdasarkan Hasil Analisis Data

Berdasarkan hasil analisis data yang telah diperoleh, tahap berikutnya adalah menyusun laporan akhir yang merangkum seluruh temuan utama serta memberikan arahan strategis untuk peningkatan mutu layanan. Laporan ini menjadi bentuk konkret dari penerjemahan data ke dalam rekomendasi yang dapat diimplementasikan. Setiap temuan dianalisis lebih lanjut untuk mengidentifikasi penyebab mendasar, sehingga solusi yang diusulkan tidak hanya bersifat reaktif tetapi juga preventif dan berkelanjutan. Selain itu, dalam tahap ini juga dilakukan penyusunan laporan akhir serta usulan strategi peningkatan kualitas layanan yang akan menjadi acuan dalam pengambilan keputusan oleh manajemen.

Contohnya, apabila ditemukan adanya ketidaktepatan waktu dalam pengambilan limbah medis, maka dapat diusulkan penggunaan sistem monitoring real-time untuk melacak jadwal pengambilan secara akurat. Di sisi lain, dalam sektor farmasi yang menunjukkan fluktuasi permintaan obat, dapat diusulkan penerapan sistem manajemen inventaris berbasis

prediksi permintaan atau integrasi sistem antara unit farmasi dengan data pelayanan medis. Setiap inisiatif diarahkan untuk menyelesaikan permasalahan yang teridentifikasi secara spesifik dan berbasis data.

Tahapan ini termasuk ke dalam fase Deployment dalam kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data mining*). Pada tahap Deployment, hasil analisis dan rekomendasi diterjemahkan ke dalam tindakan nyata yang dapat dijalankan oleh organisasi. Tujuannya adalah memastikan bahwa wawasan dan temuan dari proses analisis benar-benar diterapkan dalam operasional harian rumah sakit. Selain menyusun dokumentasi dan rencana implementasi, dilakukan pula presentasi akhir hasil pengerjaan tugas kepada tim atau pihak manajemen rumah sakit sebagai bentuk komunikasi formal terhadap temuan dan rekomendasi yang diajukan. Dengan pendekatan ini, transformasi berbasis data menjadi lebih terstruktur dan berdampak langsung terhadap peningkatan layanan.

3.3 Kendala yang Ditemukan

Selama pelaksanaan program magang di Rumah Sakit Mentari, peserta magang menghadapi sejumlah tantangan yang memengaruhi kelancaran proses kerja secara keseluruhan. Kendala tersebut tidak hanya berkaitan dengan aspek teknis, tetapi juga muncul dari dinamika komunikasi dan penyesuaian terhadap lingkungan kerja baru. Situasi ini merupakan hal yang wajar dalam proses pembelajaran di dunia profesional, khususnya dalam konteks layanan kesehatan yang kompleks. Proses adaptasi memerlukan waktu dan pemahaman yang cukup terhadap alur kerja yang ada, serta kemampuan untuk membangun koordinasi dengan berbagai pihak. Meskipun beberapa hambatan sempat memengaruhi efektivitas pelaksanaan tugas, peserta magang tetap berupaya mencari solusi dan belajar dari setiap situasi yang dihadapi. Dengan demikian, pengalaman ini menjadi bagian penting dalam proses pengembangan keterampilan praktis dan kesiapan kerja. Berikut ini adalah beberapa kendala utama yang dihadapi:

1. Akses Data yang Terbatas dan Proses Permintaan yang Bertahap

Salah satu tantangan yang cukup signifikan adalah keterbatasan akses terhadap data yang dibutuhkan. Karena data rumah sakit bersifat sensitif dan menyangkut informasi internal, peserta magang tidak diperkenankan mengakses data secara langsung. Untuk mendapatkan data, peserta harus mengajukan permintaan melalui Supervisor Divisi IT, yang kemudian diteruskan ke divisi terkait. Proses ini memerlukan waktu yang tidak sedikit, terutama jika pihak yang bersangkutan sedang memiliki beban kerja yang tinggi. Hal ini berdampak pada keterlambatan dalam memulai proses analisis data. Keterbatasan ini juga memerlukan kehati-hatian dalam menyusun permintaan data agar tepat sasaran dan tidak menambah beban komunikasi yang tidak perlu bagi pihak rumah sakit.

2. Keterbatasan Koordinasi dan Hambatan Komunikasi Teknis

Selama program magang, koordinasi antara peserta dan divisi-divisi lain seperti Farmasi, Laboratorium, dan *General Affair* terkadang mengalami keterbatasan. Padatnya jadwal kerja dari tim rumah sakit menyebabkan sulitnya menentukan waktu diskusi atau validasi data secara rutin. Selain itu, kendala komunikasi juga muncul akibat perbedaan latar belakang teknis yaitu beberapa staf di divisi non-IT masih kurang familiar dengan istilah atau konsep yang berkaitan dengan bidang teknologi informasi, seperti istilah dashboard, syntax, atau python. Hal ini membuat penyampaian maksud analisis menjadi kurang efektif dan memerlukan pendekatan komunikasi yang lebih sederhana agar tujuan diskusi dapat tersampaikan dengan baik.

3. Kualitas Data yang Belum Siap untuk Analisis Langsung

Kendala lainnya yang cukup memakan waktu adalah kondisi data yang diperoleh belum siap untuk digunakan secara langsung dalam proses analisis. Data sering kali belum lengkap, tidak terstruktur, mengandung nilai kosong, serta memiliki format yang berbeda-beda antara satu file dengan file lainnya. Sebagai contoh, terdapat kolom yang tidak konsisten dalam penamaan, atau data numerik yang masih tersimpan dalam bentuk teks. Hal ini menuntut peserta magang untuk melakukan proses pembersihan (*data*

cleaning) dan transformasi terlebih dahulu sebelum data bisa diolah lebih lanjut untuk pembuatan *dashboard* interaktif.

3.4 Solusi atas Kendala yang Ditemukan

Untuk menjawab berbagai kendala yang muncul selama pelaksanaan magang di Rumah Sakit Mentari, sejumlah solusi strategis telah diterapkan. Solusi ini bertujuan untuk memastikan bahwa kegiatan magang tetap berjalan secara efektif dan hasil akhir yang dihasilkan tetap berkualitas, meskipun dihadapkan pada berbagai keterbatasan. Pendekatan yang diambil difokuskan pada peningkatan komunikasi, fleksibilitas dalam menyusun rencana kerja, serta kemampuan adaptasi terhadap situasi yang berkembang. Selain itu, peserta magang juga berinisiatif untuk lebih proaktif dalam mencari informasi dan melakukan koordinasi lintas divisi. Adapun penjelasan detail solusi yang dilakukan adalah sebagai berikut:

1. **Koordinasi Melalui Supervisor dan Penyusunan Skala Prioritas Data**
Keterbatasan akses terhadap data internal rumah sakit menjadi tantangan utama, terutama mengingat adanya regulasi dan protokol ketat terkait keamanan dan kerahasiaan data pasien. Untuk mengatasi hal ini, peserta magang melakukan komunikasi intensif dan terstruktur dengan Supervisor dari Divisi IT Rumah Sakit Mentari. Data yang bersifat publik, administratif, atau tidak mengandung informasi pribadi diprioritaskan untuk diberikan terlebih dahulu. Dengan pendekatan ini, peserta magang tetap dapat memulai proses analisis dan pembuatan *dashboard* meskipun belum seluruh data tersedia. Komunikasi dilakukan secara bertahap untuk memastikan bahwa setiap permintaan data tidak membebani pekerjaan rutin staf IT rumah sakit.
2. **Pendekatan Komunikasi Adaptif dan Penyederhanaan Istilah Teknis**
Mengingat waktu staf rumah sakit yang terbatas, koordinasi dilakukan melalui pendekatan yang lebih fleksibel dan adaptif. Pertemuan tatap muka hanya dijadwalkan saat benar-benar diperlukan, sementara komunikasi sehari-hari dilakukan secara daring melalui WhatsApp. Pendekatan ini

memungkinkan peserta magang untuk tetap melakukan klarifikasi dan konsultasi dengan divisi terkait secara lebih cepat dan efisien, meskipun tidak dalam forum resmi. Untuk mengatasi hambatan komunikasi teknis akibat perbedaan pemahaman istilah antara peserta dan staf non-IT, peserta magang juga berinisiatif menyederhanakan istilah atau menyertakan penjelasan tambahan dalam setiap komunikasi. Selain itu, penggunaan visual sederhana seperti contoh tangkapan layar atau diagram alur turut membantu memperjelas maksud dari konsep teknologi yang disampaikan. Komunikasi yang responsif dan terbuka ini tidak hanya mempercepat proses penyampaian kendala, tetapi juga mengurangi potensi miskomunikasi serta mendukung pengambilan keputusan yang cepat dan tepat.

3. Penerapan Teknik *Data cleaning* Menggunakan Python

Untuk mengatasi permasalahan kualitas data, dilakukan proses pembersihan data (*data cleaning*) secara sistematis menggunakan Python, khususnya dengan *library* pandas. Langkah-langkah yang diambil antara lain memperbaiki format angka, dan menyesuaikan nama kolom agar lebih deskriptif dan konsisten. Setelah data dianggap bersih, data tersebut diekspor ke format Excel dan diintegrasikan ke dalam Power BI untuk dianalisis dan divisualisasikan secara interaktif. Selain itu, beberapa proses transformasi tambahan dilakukan seperti pengelompokan data berdasarkan kategori tertentu, pembuatan kolom baru agar analisis menjadi lebih mendalam dan fleksibel. Peserta magang juga mendokumentasikan setiap tahap pembersihan dan transformasi dalam bentuk *script* Python yang terdokumentasi dengan baik, agar dapat digunakan ulang dan dimodifikasi di masa mendatang. Dengan pendekatan ini, hasil analisis menjadi lebih akurat dan transparan. Penerapan teknik ini juga membantu mempercepat proses analisis karena data yang telah dibersihkan dapat langsung digunakan tanpa perlu pemeriksaan ulang secara manual. Selain itu, penggunaan Python memungkinkan otomatisasi proses *cleaning* untuk berbagai periode data secara konsisten dan efisien.