

BAB III

PELAKSANAAN KERJA MAGANG

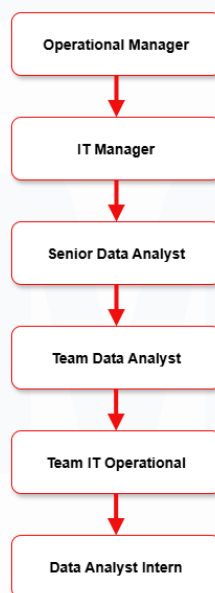
3.1 Kedudukan dan Koordinasi

Dalam pelaksanaan magang di PT Eka Satya Puspita, posisi *Data Analyst Intern* diisi oleh tim yang terdiri dari tiga mahasiswa dari berbagai universitas. Tim ini beroperasi langsung di bawah bimbingan *Data Analyst Senior* serta *IT Manager*, yang bertanggung jawab dalam memberikan tugas, arahan, serta bimbingan selama proses magang berlangsung. Struktur kerja yang terorganisir ini memastikan bahwa setiap anggota tim memahami peran dan tanggung jawabnya dalam mendukung operasional perusahaan. Selama masa magang, setiap progres aktivitas dan tugas yang diberikan wajib dilaporkan kepada *Supervisor/Senior Data Analyst*. PT Eka Satya Puspita merupakan perusahaan yang bergerak di industri logistik, di mana aktivitas operasional berjalan secara kontinu tanpa henti, termasuk di malam hari. Oleh karena itu, setiap individu dalam tim memiliki job desk tetap yang harus diselesaikan setiap hari dan diperbarui ke dalam database perusahaan. Proses pelaporan menjadi aspek krusial untuk memastikan bahwa semua tugas dapat diselesaikan dalam jangka waktu yang telah ditentukan serta memberikan umpan balik yang konstruktif guna meningkatkan keterampilan.

Koordinasi antar tim operasional merupakan aspek utama dalam menjaga standar performa perusahaan. Sebagai *Data Analyst Intern*, diwajibkan untuk berkoordinasi secara langsung dengan Tim IT Lapangan yang bertugas mengelola data dari gudang Shopee pada saat pengambilan barang. PT Eka Satya Puspita memiliki kantor di berbagai daerah atau *origin*, seperti Surabaya, Medan, Bali, dan kota-kota lainnya. Di setiap daerah ini, perusahaan juga memiliki Tim IT Lapangan yang bertanggung jawab atas pengelolaan barang *inbound* dan *outbound*. Oleh karena itu, koordinasi yang dilakukan oleh Tim *Data Analyst Intern* tidak hanya terbatas pada tim yang ada di Jakarta, tetapi juga mencakup tim dari berbagai daerah lainnya. Hal ini menjadikan komunikasi lintas daerah sebagai faktor penting dalam memastikan akurasi dan kelancaran aliran data logistik. Sebagai bagian dari Tim IT

Intern, koordinasi terkait permintaan data (*internal/external data request*) serta pengumpulan data (*data collection*) dilakukan secara rutin guna diperbarui ke dalam database perusahaan. Data ini kemudian digunakan untuk proses pemantauan logistik secara *real-time* demi menjaga *SLA* (*Service Level Agreement*). Setelah barang sampai di tujuan, data tersebut diolah lebih lanjut oleh Tim *Data Analyst* untuk dianalisis sebagai bahan evaluasi performa logistik.

Dengan demikian, kedudukan dan koordinasi yang terjalin selama program magang menjadi salah satu fondasi utama dalam meningkatkan kualitas operasional PT Eka Satya Puspita. Keberhasilan kolaborasi antara *Data Analyst Intern*, Tim IT Operasional, serta Tim Manajemen tidak hanya meningkatkan pengalaman belajar para *intern*, tetapi juga berkontribusi dalam menjaga integritas dan efisiensi operasional perusahaan. Untuk memberikan gambaran lebih jelas mengenai struktur kedudukan dalam konteks tim, berikut disajikan Gambar 3.1 yang menyoroti hubungan antara posisi magang dengan para senior serta jalur koordinasi yang terjadi selama program magang berlangsung.

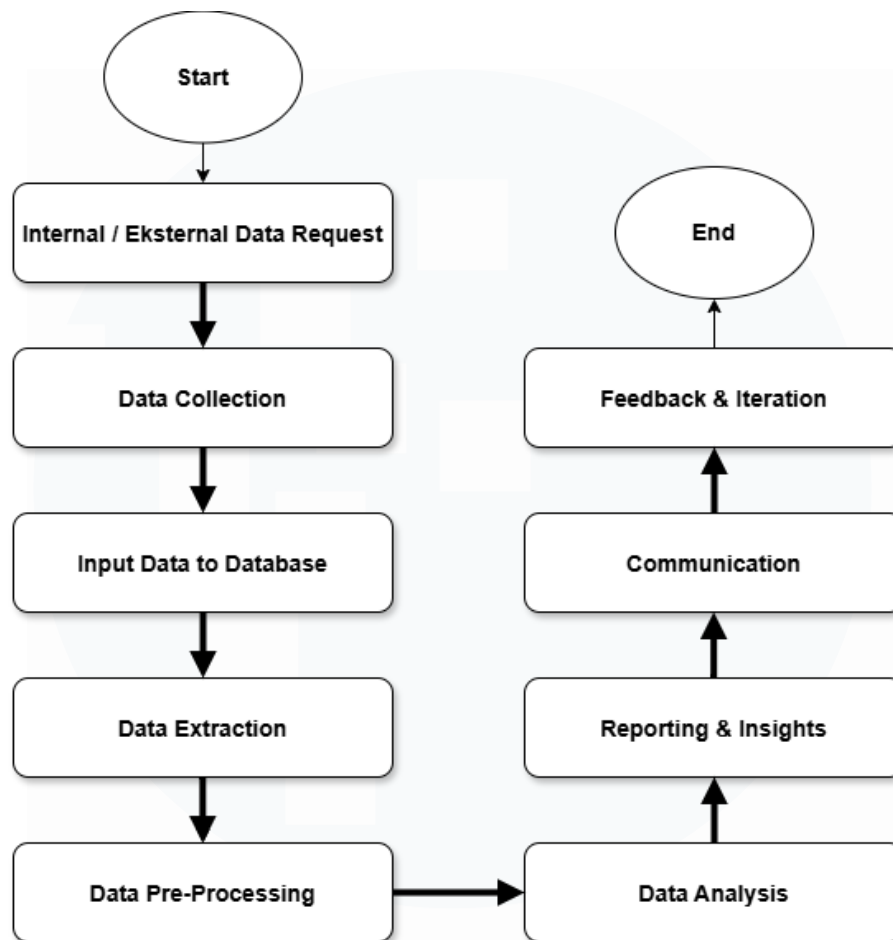


Gambar 3. 1 Diagram Hierarki Posisi Data Analyst Intern

Gambar 3.1 Struktur koordinasi dalam tim *Data Analyst Intern* di PT Eka Satya Puspita dirancang untuk memastikan efisiensi dalam pengelolaan data

logistik. Setiap tingkatan dalam hierarki memiliki peran yang saling terhubung, memungkinkan aliran informasi dan pengambilan keputusan yang lebih cepat dan akurat. *Operational Manager* memiliki tanggung jawab utama dalam memastikan operasional berjalan sesuai standar, dengan *IT Manager* sebagai penghubung utama yang mengelola sistem informasi dan infrastruktur teknologi yang digunakan dalam operasional logistik. Pada level berikutnya, *Senior Data Analyst* bertanggung jawab dalam pengolahan dan analisis data yang lebih mendalam. Peran ini juga mencakup koordinasi dan supervisi terhadap Tim *Data Analyst*, yang berfokus pada pemantauan tren operasional serta penyajian data dalam bentuk yang lebih mudah dipahami oleh manajemen. Tim ini bekerja berdampingan dengan Tim *IT Operasional*, yang menangani pencatatan dan verifikasi data barang masuk serta keluar di berbagai titik distribusi yang tersebar di seluruh Indonesia.

Berdasarkan struktur ini, *Data Analyst Intern* ditempatkan sebagai bagian dari tim yang bertanggung jawab untuk membantu proses pengumpulan, pemrosesan, dan analisis data. Untuk memastikan kualitas data tetap terjaga, koordinasi dilakukan secara intensif dengan Tim IT Operasional yang bertugas di berbagai kantor cabang perusahaan di berbagai daerah seperti Jakarta, Surabaya, Medan, dan Bali. Interaksi antara tim memastikan bahwa data yang digunakan dalam analisis telah melalui tahap validasi yang ketat sebelum akhirnya diproses lebih lanjut menjadi informasi strategis bagi manajemen. Dengan adanya sistem koordinasi yang terstruktur ini, setiap informasi yang dihasilkan dapat digunakan untuk mendukung pengambilan keputusan yang lebih akurat dan responsif terhadap perubahan di lapangan. Selain itu, kolaborasi lintas tim ini juga mendorong peningkatan standar pelaporan serta konsistensi format data antar cabang. Hal ini penting untuk menjaga kesinambungan proses analisis dalam jangka panjang. Koordinasi semacam ini juga berperan penting dalam mencegah miskomunikasi data yang dapat berdampak pada ketidaktepatan analisis atau interpretasi hasil. Untuk memahami bagaimana setiap tahap dalam pengolahan data dilakukan dari awal hingga akhir, *Flow Chart Alur Kerja* yang disajikan dalam Gambar 3.2 berikut akan memberikan gambaran lebih jelas mengenai proses tersebut.



Gambar 3. 2 Flow Chart Alur Kerja Magang

Gambar 3.2 setiap langkah dalam alur kerja ini memiliki peran krusial dalam memastikan bahwa data yang dihasilkan memenuhi standar kualitas serta kebutuhan operasional, baik untuk data *internal* maupun *external*. Proses ini diawali dengan tahap Start, yang menjadi titik awal dalam siklus pengolahan data. Pemicu utama dalam tahap ini adalah adanya kebutuhan informasi yang kemudian dituangkan dalam bentuk *Internal / External Data Request* atau Permintaan Data *Internal / External*. Permintaan ini berasal dari berbagai divisi *internal* perusahaan, seperti tim operasional yang membutuhkan data terkait pengiriman barang, manajemen yang memerlukan laporan kinerja logistik, atau tim keuangan yang memerlukan data transaksi terkait ongkos kirim dan biaya operasional. Selain itu, permintaan juga dapat datang dari pihak eksternal seperti mitra bisnis yang membutuhkan informasi spesifik terkait status pengiriman, jumlah barang yang

diterima, atau kepastian jadwal distribusi. Permintaan ini harus dibuat secara jelas dan terstruktur agar dapat diproses dengan cepat dan akurat, sehingga tidak menimbulkan kesalahan dalam tahap selanjutnya.

Setelah permintaan data diterima, tahap berikutnya adalah *Data Collection* atau Pengumpulan Data, di mana informasi dikumpulkan dari berbagai sumber yang relevan. Sumber utama dalam pengumpulan data ini adalah Tim *IT Operational* yang bertanggung jawab dalam mengelola dan memonitor pergerakan barang inbound maupun outbound. Data yang dikumpulkan meliputi detail pengiriman seperti jumlah dan jenis barang yang dikirim, destinasi tujuan, berat kargo, serta informasi penerbangan seperti jenis maskapai yang digunakan, estimasi waktu keberangkatan dan kedatangan (ETA & ATA), serta estimasi dan aktual waktu pengiriman dari gudang (ETD & ATD). Selain itu, Status Kedatangan *Vendor* saat barang tiba juga menjadi bagian penting dalam pengumpulan data untuk memastikan kelancaran proses logistik, dengan indikator seperti (*On Time*) jika barang tiba sesuai jadwal atau (*Late*) jika terjadi keterlambatan dalam pengiriman. Ketelitian dalam tahap ini sangat diperlukan, karena data yang tidak lengkap atau tidak akurat dapat menghambat proses selanjutnya dan berpotensi menyebabkan kesalahan dalam analisis serta pengambilan keputusan.

Setelah data terkumpul, tahap *Input Data to Database* dilakukan untuk menyimpan serta mengorganisir data ke dalam sistem PT Eka Satya Puspita. Penyimpanan data dilakukan secara sistematis agar mudah diakses dan digunakan kembali ketika diperlukan untuk keperluan analisis atau pelaporan. Data yang telah diinput ke dalam database dipantau secara berkala untuk memastikan bahwa semua informasi telah terekam dengan benar serta sesuai dengan kondisi aktual di lapangan. Jika ditemukan ketidaksesuaian, maka akan dilakukan validasi ulang sebelum melanjutkan ke tahap berikutnya. Setelah data tersimpan dengan baik, proses *Data Extraction* dilakukan untuk menarik data dari sistem guna diproses lebih lanjut. Tahap ekstraksi ini bertujuan untuk memastikan bahwa hanya data yang relevan dan diperlukan yang diambil, sehingga tidak ada informasi berlebihan yang dapat memperlambat analisis.

Tahap berikutnya adalah *Data Pre-Processing*, yang merupakan bagian penting dalam memastikan kualitas data sebelum dianalisis. Di PT Eka Satya Puspita, proses ini dilakukan secara berkala setiap sepuluh hari sekali, mengikuti sistem penyimpanan yang dibagi ke dalam tiga periode dalam satu bulan, yaitu tanggal 1-10, 11-20, dan 21-31. Pada tahap ini, dilakukan berbagai langkah pemrosesan data seperti data *cleansing* untuk menghilangkan informasi yang tidak valid, data formatting agar format data seragam, serta penghapusan data duplikat guna mencegah kesalahan analisis. Selain itu, jika terdapat data yang tidak lengkap, dilakukan pengecekan lebih lanjut untuk mengisi kekosongan tersebut agar hasil analisis dapat lebih akurat. Tahap ini sangat krusial karena kesalahan dalam *pre-processing* dapat menyebabkan misinterpretasi data dalam tahap analisis berikutnya.

Setelah data diproses dengan baik, tahap *Data Analysis* dilakukan untuk mengolah informasi menjadi wawasan yang berguna dalam operasional perusahaan. Analisis dilakukan menggunakan teknik seperti *pivot table* dan visualisasi data dalam Excel, yang membantu dalam mengidentifikasi pola, tren, serta potensi kendala yang dapat terjadi dalam proses logistik. Analisis ini dapat memberikan berbagai *insight*, seperti efisiensi waktu pengiriman, tingkat keterlambatan barang, performa vendor, serta perbandingan volume pengiriman dari bulan ke bulan. Hasil dari tahap analisis ini menjadi dasar bagi pengambilan keputusan strategis yang dapat membantu dalam meningkatkan efisiensi operasional serta mengoptimalkan kinerja perusahaan. Setelah hasil analisis diperoleh, tahap Reporting & Insights dilakukan untuk merangkum temuan utama ke dalam bentuk laporan yang lebih mudah dipahami. Laporan ini dikirimkan kepada Senior Data Analyst untuk ditinjau dan divalidasi sebelum didistribusikan lebih lanjut ke manajemen atau divisi terkait lainnya. Selain sebagai bahan evaluasi internal, laporan ini juga digunakan dalam rapat mingguan, di mana berbagai pemangku kepentingan dapat membahas langkah-langkah perbaikan yang diperlukan berdasarkan data yang telah dianalisis. Laporan ini berperan sebagai jembatan antara hasil analisis dan tindak lanjut di lapangan.

Tahap berikutnya adalah *Communication*, yang berperan penting dalam memastikan bahwa informasi yang telah diproses dan dianalisis dapat dipahami serta digunakan oleh pihak-pihak yang berkepentingan. Hasil analisis yang telah dirangkum dalam laporan disampaikan kepada tim *internal* terkait, seperti manajer IT, tim operasional, serta divisi manajemen agar semua pihak memiliki pemahaman yang sama mengenai kondisi terkini perusahaan. Komunikasi yang baik dalam tahap ini sangat diperlukan, karena keputusan strategis yang diambil bergantung pada pemahaman yang tepat terhadap data yang telah disajikan. Jika terjadi miskomunikasi atau interpretasi yang salah, maka keputusan yang diambil dapat menjadi tidak efektif atau bahkan merugikan perusahaan.

Sebagai langkah terakhir, tahap *Feedback & Iteration* dilakukan sebagai bentuk evaluasi terhadap seluruh proses yang telah berjalan. Pada tahap ini, berbagai pihak yang terlibat dalam siklus pengolahan data memberikan umpan balik terkait keakuratan serta efektivitas data yang telah dianalisis. Apabila terdapat kekeliruan atau kekurangan dalam hasil analisis, dilakukan penyesuaian ulang terhadap metode dan pelaporan guna meningkatkan efisiensi serta menjamin keakuratan dan kredibilitas informasi yang dihasilkan. Selain itu, *feedback* juga digunakan untuk mengidentifikasi area yang dapat diperbaiki atau dioptimalkan dalam proses pengelolaan data di masa mendatang.

Dengan adanya alur kerja yang sistematis serta proses yang terstruktur dari tahap permintaan hingga evaluasi, pengelolaan data di PT Eka Satya Puspita dapat berjalan lebih akurat, efisien, dan dapat diandalkan sebagai dasar dalam pengambilan keputusan operasional. Hal ini juga memastikan bahwa perusahaan dapat terus meningkatkan kualitas layanannya serta merespons dinamika bisnis dengan lebih baik berdasarkan analisis berbasis data yang komprehensif.

3.2 Tugas dan Uraian Kerja Magang

Sebagai *intern* di posisi *Data Analyst* di PT Eka Satya Puspita, peran utama selama pelaksanaan magang adalah melakukan pengolahan informasi berbasis data guna menunjang kegiatan operasional perusahaan. Tugas yang dijalankan meliputi proses pengumpulan, pembersihan, analisis, serta penyusunan laporan dari data

logistik yang diperoleh melalui berbagai sumber internal maupun mitra eksternal. Dalam praktiknya, data yang telah dikumpulkan diproses lebih lanjut untuk memastikan keakuratannya sebelum dianalisis guna mendapatkan wawasan yang berguna bagi evaluasi dan pengambilan keputusan. Selain analisis operasional, perbandingan dengan kompetitor juga menjadi bagian dari tugas yang dijalankan. Informasi terkait strategi, layanan, dan performa pesaing dikumpulkan dan diolah untuk memahami posisi perusahaan di industri. Hasil dari analisis ini kemudian disusun dalam laporan yang disampaikan kepada tim terkait melalui dokumen dan presentasi dalam pertemuan rutin. Di samping itu, pemantauan serta evaluasi kinerja turut menjadi bagian dari tanggung jawab selama magang. Hal ini dilakukan dengan mengembangkan indikator kinerja utama (KPI) yang membantu dalam menilai efektivitas strategi perusahaan. Koordinasi dengan tim *Data Analyst* serta departemen lain juga menjadi aspek penting untuk memastikan bahwa seluruh proses pengolahan data berjalan sesuai kebutuhan bisnis. Dengan peran ini, keterampilan analitis serta komunikasi menjadi aspek yang sangat ditekankan dalam mendukung kelancaran operasional berbasis data. Maka, berdasarkan seluruh tugas yang telah dijalankan, Tabel 3.1 menyajikan realisasi dari agenda kerja yang telah dirumuskan sebelumnya.

Tabel 3. 1 Tabel Realisasi Agenda Data Quality Intern

No	Kegiatan	Tanggal Mulai	Tanggal Selesai
1	Memahami Konteks Bisnis dan Tujuan Analisis (<i>Business Understanding</i>)		
1.1	Memahami Proses bisnis dan alur distribusi kargo di PT Eka Satya Puspita	10 Februari 2025	10 Maret 2025
1.2	Analisis kompetitor (CKL Cargo dan lainnya)	22 Mei 2025	3 Juni 205
2	Eksplorasi dan Validasi Sumber Data (<i>Data Understanding</i>)		
2.1	Mengumpulkan data dari berbagai sumber <i>internal</i> dan eksternal	10 Februari 2025	25 Juni 2025
2.2	Memastikan data yang dikumpulkan relevan dan akurat	10 Februari 2025	25 Juni 2025
3	Persiapan Data untuk Analisis Lanjutan (<i>Data Preparation</i>)		
3.1	Aplikasi analisis data (Excel, & Python)	10 Februari 2025	25 Juni 2025

No	Kegiatan	Tanggal Mulai	Tanggal Selesai
3.2	Melakukan <i>data cleaning</i>	10 Februari 2025	25 Juni 2025
3.3	Mengolah data ke format yang siap dianalisis	10 Februari 2025	25 Juni 2025
4	Penerapan Teknik Analisis		
4.1	Segmentasi atau klasifikasi awal (<i>K-Means</i>)	25 April 2025	15 Mei 2025
5	Penilaian Kinerja Model (<i>Evaluation</i>)		
5.2	Evaluasi Kinerja Model	20 Mei 2025	25 Mei 2025
6	Implementasi Hasil (<i>Deployment</i>)		
6.1	Pembuatan & Implementasi <i>Dashboard</i> Interaktif	27 Mei 2025	20 Juni 2025

Table 3.1 merujuk pada pelaksanaan program magang di PT Eka Satya Puspita, pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*) digunakan sebagai kerangka kerja utama dalam merancang dan menjalankan agenda kerja. Pendekatan ini dipilih untuk memastikan bahwa seluruh proses analisis data dilakukan secara sistematis, terarah, dan mampu menghasilkan *insight* yang bermanfaat bagi peningkatan kinerja operasional perusahaan. Tabel 3.1 merinci agenda kerja yang telah direalisasikan sesuai dengan enam tahap utama dalam CRISP-DM. Pada tahap *Business Understanding*, fokus utama diarahkan untuk memahami secara komprehensif proses bisnis dan alur distribusi kargo yang dijalankan oleh PT Eka Satya Puspita. Hal ini dilakukan agar setiap langkah analisis yang dilakukan nantinya benar-benar relevan dan selaras dengan kebutuhan operasional perusahaan. Selain itu, dilakukan juga analisis terhadap kompetitor seperti CKL Cargo dan lainnya, yang kemudian dirangkum dalam bentuk laporan ringkas guna memberi gambaran mengenai posisi kompetitif perusahaan dalam industri logistik. Tahap *Data Understanding* mencakup kegiatan pengumpulan data dari berbagai sumber *internal* maupun *eksternal* yang relevan, seperti data pengiriman, keterlambatan, rute, dan sebagainya. Fokus dalam tahap ini adalah memastikan bahwa data yang diperoleh memiliki kualitas yang baik, valid, serta mencerminkan realitas operasional perusahaan, sehingga dapat digunakan secara

optimal dalam analisis lanjutan. Setelah data terkumpul, tahapan selanjutnya adalah *Data Preparation*, yang meliputi proses *data cleaning* untuk menghilangkan duplikasi atau inkonsistensi, serta konversi data ke dalam format yang siap dianalisis. Proses ini menjadi fondasi penting sebelum dilakukan analisis lanjutan, karena kualitas hasil analisis sangat dipengaruhi oleh kebersihan dan kesiapan data. Pada tahap Modeling, dilakukan berbagai teknik analisis menggunakan alat bantu seperti *Microsoft Excel dan Python*. Salah satu pendekatan analisis yang direncanakan adalah penerapan teknik *K-Means Clustering* untuk mengelompokkan rute pengiriman berdasarkan keterlambatan atau efisiensi waktu tempuh. Selain itu, dibuat pula *dashboard* interaktif untuk memvisualisasikan data dan *insight* yang dihasilkan. *Dashboard* ini ditujukan sebagai alat bantu pengambilan keputusan bagi manajemen dalam memahami performa operasional secara berkala dan lebih intuitif.

Tahap *Evaluation* difokuskan pada evaluasi hasil analisis dan performa model atau *dashboard* yang dikembangkan. Selain itu, dilakukan juga uji coba *dashboard* bersama *supervisor* untuk mendapatkan umpan balik serta memastikan *dashboard* dapat digunakan secara fungsional dan efektif. Terakhir, pada tahap Deployment, hasil dari keseluruhan proses dianalisis dan disusun dalam bentuk laporan yang kemudian dipresentasikan kepada tim atau stakeholder terkait. Selain menyampaikan temuan analisis, tahap ini juga melibatkan kolaborasi lintas tim, terutama dengan tim *Data Analyst* dan divisi operasional, guna memastikan bahwa *insight* yang dihasilkan dapat diimplementasikan secara nyata dalam mendukung pengambilan keputusan dan peningkatan strategi distribusi perusahaan. Melalui pemetaan agenda kerja yang sistematis berdasarkan CRISP-DM ini, program magang tidak hanya berkontribusi secara teknis melalui visualisasi data dan analisis statistik, tetapi juga membawa dampak strategis terhadap efisiensi dan efektivitas operasional PT Eka Satya Puspita. Pada sub-bab selanjutnya, pembahasan akan difokuskan pada tugas dan uraian kerja magang secara lebih rinci, termasuk kontribusi langsung yang diberikan terhadap proyek atau inisiatif perusahaan selama masa penugasan berlangsung.

3.2.1 Memahami Konteks Bisnis & Tujuan Analisis (*Business Understanding*)

Pada tahap Business Understanding, langkah awal yang dilakukan adalah memahami secara menyeluruh konteks bisnis dari PT Eka Satya Puspita (ESP) sebagai perusahaan yang bergerak di bidang distribusi kargo udara. Pemahaman ini menjadi pondasi penting untuk memastikan bahwa proses analisis data yang dijalankan selaras dengan kebutuhan dan tujuan bisnis perusahaan. Melalui tahap ini, peneliti berupaya menggali informasi terkait alur operasional pengiriman barang, tantangan yang dihadapi dalam distribusi, serta target peningkatan efisiensi yang ingin dicapai. Selain itu, dalam proses ini juga ditetapkan tujuan analisis, yaitu untuk mengidentifikasi peluang perbaikan dalam distribusi, memetakan rute pengiriman yang lebih optimal, serta menyusun rekomendasi yang dapat mendukung pengambilan keputusan strategis oleh manajemen. Dengan landasan pemahaman yang kuat terhadap proses bisnis dan tujuan analisis, tahap-tahap selanjutnya dapat dijalankan secara lebih terarah dan efektif.

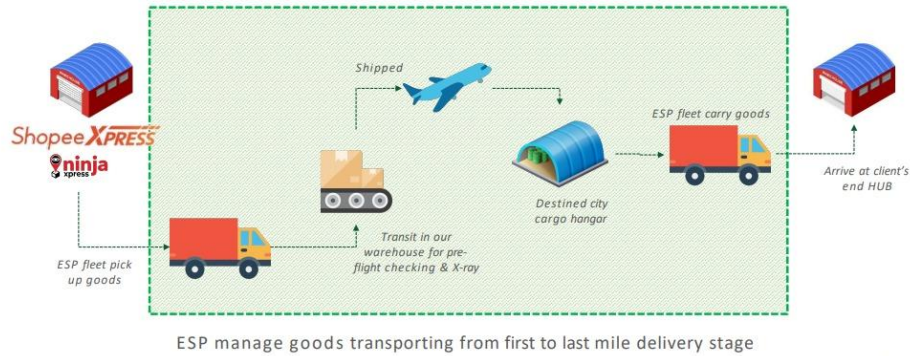
1. Memahami Proses bisnis dan alur distribusi kargo di PT Eka Satya Puspita

Langkah awal dalam memahami konteks bisnis PT Eka Satya Puspita adalah dengan memetakan secara rinci proses distribusi kargo yang dijalankan perusahaan. Sebagai perusahaan yang bergerak di bidang logistik udara, PT Eka Satya Puspita memiliki peranan penting dalam menghubungkan rantai pasok dari gudang asal milik klien hingga gudang tujuan yang tersebar di berbagai kota. Memahami alur distribusi ini menjadi landasan penting dalam proses analisis, terutama untuk mengidentifikasi potensi hambatan, ketidakefisienan, maupun peluang optimasi di setiap tahap pengiriman. Untuk memberikan gambaran yang lebih jelas, alur distribusi kargo di PT Eka Satya Puspita dapat dilihat pada Gambar 3.3 berikut, yang menyajikan siklus penuh manajemen kargo dari awal hingga akhir proses pengiriman.

FULL CYCLE OF OUR CARGO MANAGEMENT



We basically connect the supply chain from client's originating warehouse to destined warehouse



Gambar 3.3 Full Cycle of PT Eka Satya Puspita Cargo Management

(Sumber: Divisi *Data Analyst* PT Eka Satya Puspita)

Gambar 3,3 proses distribusi kargo di PT Eka Satya Puspita berjalan melalui serangkaian tahapan yang kompleks dan terstruktur, yang bertujuan untuk memastikan kelancaran pengiriman dari titik asal hingga ke tujuan akhir. Sebagai vendor logistik yang bermitra dengan Shopee Express, PT Eka Satya Puspita berperan penting dalam pengiriman barang antar pulau, khususnya ke wilayah Sumatera dan Jawa. Alur distribusi dimulai dari pengambilan barang oleh armada PT Eka Satya Puspita di berbagai *Distribution Center (DC)* besar milik Shopee, seperti DC Kapuk, DC Sunter, dan DC Cakung. Barang-barang yang diterima ini berasal dari berbagai daerah di Indonesia, bahkan dari luar negeri seperti China, yang masuk ke Indonesia melalui jalur laut maupun udara. Setibanya di DC Shopee, barang-barang tersebut terlebih dahulu menjalani proses pensortiran berdasarkan destinasi pulau dan kota tujuan. DC Kapuk umumnya menangani barang dengan tujuan Pulau Sumatera, DC Sunter melayani wilayah Pulau Jawa, sedangkan DC Cakung mencakup area Jabodetabek dan sekitarnya. Setelah proses penyortiran selesai dan barang sudah dikelompokkan per kota tujuan, PT Eka Satya Puspita mengambil alih peran distribusi. Barang diangkut menggunakan armada truk jenis CDD Long oleh tim lapangan dan pengemudi PT Eka Satya Puspita, lalu dikirim

menuju gudang utama perusahaan yang berlokasi di kawasan Kapuk Kemayoran, Jakarta.

Dalam gudang utama ini, barang-barang yang masuk melalui proses *unloading* dan disiapkan untuk tahap berikutnya, yaitu proses reservasi slot penerbangan. Tim reservasi PT Eka Satya Puspita melakukan penjadwalan untuk memastikan setiap karung barang mendapatkan slot di maskapai tujuan sesuai dengan destinasi masing-masing. Selanjutnya, setiap karung barang akan diberi label berupa barcode *Air Way Bill (AWB)* yang berfungsi sebagai tanda pengenal dan tiket masuk ke fasilitas kargo maskapai di bandara. Setelah proses labeling selesai, barang-barang ini dimuat kembali ke dalam truk milik PT Eka Satya Puspita untuk dikirim menuju gudang kedua, yaitu gudang imigrasi barang yang berfungsi sebagai tempat pemeriksaan pra-penerbangan.

Setiap paket wajib melalui proses *pre-flight checking* dan *X-ray screening* untuk memastikan tidak ada barang berbahaya yang masuk dalam muatan pesawat. Proses ini menjadi krusial dalam menjamin keselamatan penerbangan, dengan memastikan semua paket terbebas dari *Dangerous Goods (DG)* seperti bahan peledak, cairan mudah terbakar, zat beracun, atau material radioaktif. Jika ditemukan barang yang terindikasi DG, maka barang tersebut akan dibuka untuk pemeriksaan lebih lanjut. Apabila terbukti sebagai barang berbahaya, paket tersebut akan dikembalikan ke gudang utama Shopee dan dialihkan untuk pengiriman jalur darat yang sesuai dengan prosedur. Setelah seluruh barang dinyatakan lolos dari proses *X-ray*, maka muatan siap untuk melanjutkan perjalanan menuju terminal kargo di bandara sebagai tahap selanjutnya dalam siklus distribusi.

Setelah barang-barang dari PT Eka Satya Puspita berhasil melewati proses pemeriksaan di gudang imigrasi barang, perjalanan logistik berlanjut ke tahap berikutnya, yakni memasuki terminal kargo bandara. Di sini, barang kembali melalui proses *unloading*, lalu disortir lebih lanjut berdasarkan maskapai penerbangan yang akan mengangkut barang tersebut

dan juga slot waktu keberangkatan masing-masing pesawat. Penyortiran ini menjadi krusial untuk memastikan setiap paket dapat termuat sesuai dengan jadwal pesawat yang sebelumnya telah dipesan dalam proses reservasi slot maskapai. Ketepatan proses ini mempengaruhi kelancaran distribusi di tahap-tahap selanjutnya, terutama dalam menjaga ketepatan waktu pengiriman lintas kota dan pulau. Setelah proses penyortiran selesai, barang-barang menunggu giliran untuk memasuki tahap loading ke dalam bagasi pesawat.

Proses ini tidak sekadar memindahkan barang, namun juga melibatkan penyesuaian dengan kapasitas kargo pesawat, peraturan keselamatan penerbangan, dan standar keamanan bandara. Barang-barang ini kemudian diterbangkan menuju destinasi kota tujuan masing-masing, sesuai dengan rencana pengiriman yang telah disusun oleh tim logistik PT Eka Satya Puspita. Menariknya, selama proses penerbangan ini berlangsung, seluruh alur pengiriman dipantau secara berkala oleh tim monitoring *internal* perusahaan. Untuk memastikan akurasi dan kecepatan informasi, tim ini menggunakan aplikasi *Flightradar24*, yang memungkinkan mereka melacak posisi pesawat, memperkirakan waktu keberangkatan (*Estimated Time of Departure/ETD*), serta memantau waktu kedatangan aktual (*Estimated Time of Arrival/ETA*).

Sistem monitoring ini memiliki peran vital dalam menjaga keandalan layanan PT Eka Satya Puspita. Jika terjadi *delay* atau perubahan jadwal penerbangan yang di luar perkiraan, tim dapat segera mengambil langkah antisipatif, seperti mengoordinasikan ulang jadwal penjemputan atau memberi informasi lebih cepat kepada tim di kota tujuan. Dengan begitu, potensi keterlambatan dalam proses distribusi dapat diminimalisasi, menjaga kepuasan pelanggan tetap tinggi. Setibanya di bandara tujuan, barang-barang kembali melalui proses unloading dan langsung dipindahkan ke truk kargo PT Eka Satya Puspita yang sebelumnya sudah disiapkan dan standby di terminal kargo. Proses pemindahan ini berlangsung cepat, agar

barang segera dibawa menuju gudang pensortiran akhir di kota destinasi. Di gudang ini, barang-barang disortir kembali secara lebih spesifik, yakni berdasarkan wilayah pengantaran yang lebih kecil seperti kecamatan atau kelurahan. Lokasi-lokasi penyortiran ini dikenal sebagai mini-hub, yang biasanya berupa ruko atau gudang kecil yang tersebar di berbagai wilayah dalam kota. Dari mini-hub inilah barang akhirnya masuk ke tahap *last mile delivery*, yaitu pengantaran langsung ke alamat rumah atau kantor konsumen.

Keseluruhan proses yang panjang dan berlapis ini mulai dari pengambilan barang di *Distribution Center* (DC) Shopee, penyortiran, pengiriman antar pulau, hingga pengantaran ke mini-hub semuanya dipantau secara ketat dan didokumentasikan dalam sistem database *internal* milik PT Eka Satya Puspita. Setiap tahapan proses dicatat, mulai dari jumlah paket yang diproses, waktu pengiriman, hingga data terkait keterlambatan atau kendala lain di lapangan. Data ini kemudian dikumpulkan sebagai bagian dari rekap performa bulanan perusahaan. Dengan memiliki data yang *reliable*, lengkap, dan terstruktur ini, PT Eka Satya Puspita mampu melakukan analisis lanjutan untuk mengevaluasi efektivitas operasional dan merancang strategi distribusi yang lebih efisien di masa depan. Setelah memahami secara menyeluruh alur distribusi kargo yang dijalankan oleh PT Eka Satya Puspita, tahap berikutnya dalam proses analisis adalah melakukan pemetaan dan kajian terhadap kompetitor. Langkah ini penting untuk mengetahui posisi perusahaan di tengah persaingan industri logistik nasional, serta untuk mengidentifikasi peluang dan tantangan yang dapat dioptimalkan ke depannya.

2. Analisis kompetitor (CKL Cargo dan lainnya)

Dalam tahap *Business Understanding*, salah satu agenda penting yang dilakukan adalah melakukan analisis kompetitor, khususnya terhadap perusahaan yang bergerak di segmen layanan logistik serupa dengan PT Eka Satya Puspita. Dua nama yang cukup menonjol di segmen ini adalah CKL

Cargo dan PT Inti Duta Logistik, yang keduanya aktif menjadi vendor pengiriman barang dalam jaringan *e-commerce* besar seperti Shopee *Express*. Melalui analisis kompetitor ini, diharapkan PT Eka Satya Puspita dapat memahami posisi dan daya saingnya di pasar logistik *middle-mile* maupun *freight forwarding*, sekaligus mengidentifikasi potensi peluang dan ancaman yang dihadapi. Untuk memberikan gambaran yang lebih jelas mengenai posisi strategis PT Eka Satya Puspita dibandingkan dengan para kompetitornya, berikut disajikan Tabel 3.2 yang memuat perbandingan analisis SWOT antara PT Eka Satya Puspita, CKL Cargo, dan Intiduta Logistic.

Tabel 3. 2 Perbandingan SWOT antara PT ESP, CKL Cargo, dan Intiduta Logistik

(Sumber: Website CKL Cargo [11] dan Intiduta Logistik [12])

Aspek	PT Eka Satya Puspita	CKL Cargo	Intiduta Logistik
<i>Strengths</i> (Kekuatan)	- Mitra resmi Shopee Express untuk jalur udara ke Sumatera & Jawa - Pengelolaan proses cargo terintegrasi (reservasi, X-ray, monitoring) - Tim monitoring aktif dengan <i>live tracking</i> pesawat	- Jaringan distribusi nasional yang luas - Sistem pelacakan real-time - Kerja sama strategis dengan maskapai & kurir nasional	- Reputasi kuat di jalur Batam-Singapura - Layanan pelanggan intensif (24/7 support) - Layanan LCL/FCL dan FTL/LTL lengkap
<i>Weaknesses</i> (Kelemahan)	- Skala operasi masih dalam tahap pengembangan - Jaringan distribusi luar Jawa-Sumatera terbatas - Perlu akselerasi dalam transformasi digital	- Keterbatasan dalam jangkauan <i>last-mile delivery</i> - Tekanan harga dari ekspansi logistik in-house e-commerce	- Skala nasional masih terbatas - Investasi teknologi belum seagresif pemain besar

Aspek	PT Eka Satya Puspita	CKL Cargo	Intiduta Logistik
<i>Opportunities</i> (Peluang)	- Ekspansi jalur distribusi baru ke Kalimantan & Sulawesi - Peningkatan layanan berkala <i>monitoring</i> - Pengembangan <i>dashboard</i> analitik untuk pengambilan keputusan	- Diversifikasi layanan ke logistik cross-border - Penguatan sistem digitalisasi rute dan pelacakan	- Ekspansi dari Batam ke wilayah Sumatera besar - Penawaran paket layanan khusus untuk e-commerce
<i>Threats</i> (Ancaman)	- Persaingan dari vendor lain di Shopee Express - Dominasi pemain besar seperti SiCepat, J&T Cargo - Tekanan margin dari pengetatan tarif logistik	- Tekanan dari pemain besar logistik nasional - Perubahan strategi logistik oleh platform e-commerce	- Dominasi vendor besar di jalur utama Sumatera-Jawa - Kenaikan biaya operasional jalur udara & laut

Table 3.2 menunjukkan CKL Cargo dikenal sebagai perusahaan logistik yang menawarkan layanan pengiriman barang melalui jalur darat, laut, dan udara dengan mengusung nilai efisiensi biaya dan kecepatan waktu tempuh. Sejak didirikan pada tahun 2015, CKL telah berkembang menjadi salah satu pemain menengah yang cukup solid dengan jaringan distribusi domestik yang luas serta kemitraan strategis bersama berbagai maskapai seperti Garuda dan Lion Air, serta perusahaan kurir seperti JNE, Ninja Xpress, hingga Shopee Xpress sendiri. Keunggulan utama CKL terletak pada kemampuan mereka dalam mengintegrasikan sistem teknologi pelacakan secara berkala dan pengelolaan rute distribusi yang efisien. Meskipun begitu, skala operasi CKL masih berada di bawah pemain besar nasional, sehingga mereka tetap menghadapi keterbatasan dalam hal jangkauan last-mile delivery dan sumber daya skala besar. Selain itu,

tekanan harga yang semakin ketat akibat ekspansi logistik in-house oleh platform besar seperti Shopee Xpress (yang kini menangani sekitar 40% volume Shopee) menjadi tantangan serius bagi CKL dalam mempertahankan margin dan volume pengiriman. Dalam konteks persaingan, CKL bersinggungan langsung dengan PT Eka Satya Puspita, khususnya dalam layanan pengangkutan barang skala grosir dan pengelolaan distribusi ke wilayah Sumatera dan Jawa.

Sebagai perbandingan PT Inti Duta Logistik (Intiduta) yang berbasis di Batam menjadi kompetitor lain yang cukup relevan. Didirikan pada tahun 2008, Intiduta telah membangun reputasi kuat di kalangan pelanggan regional dengan layanan pengiriman yang mencakup jalur darat, laut, dan udara. Intiduta menawarkan berbagai layanan seperti pengiriman laut LCL/FCL, angkutan darat FTL/LTL, dan ekspedisi cepat antar kota yang dipadukan dengan layanan pelanggan intensif, termasuk komunikasi 24/7 dan sistem pelacakan berbasis web. Kelebihan Intiduta terletak pada fokus mereka di jalur lintas Batam-Singapura serta kecepatan layanan di wilayah Sumatera dan Kalimantan, menjadikannya salah satu pilihan vendor utama untuk *e-commerce* yang membutuhkan pengiriman efisien dari Batam ke kota-kota besar di Indonesia. Namun, seperti CKL, Intiduta juga menghadapi tantangan dari dominasi pemain logistik besar nasional dan tekanan harga dari strategi logistik *in-house e-commerce* yang masif. Dengan skala yang lebih menengah, Intiduta masih menghadapi keterbatasan dari sisi investasi teknologi dan perluasan jaringan nasional.

Melalui analisis terhadap dua kompetitor utama ini, PT Eka Satya Puspita dapat mengidentifikasi keunggulan yang dapat terus diperkuat, seperti layanan pergudangan dan pengelolaan kargo terintegrasi yang menjadi kekuatan unik dalam jaringan Shopee Express. Selain itu, analisis ini juga memberikan gambaran bahwa untuk tetap kompetitif di tengah dinamika pasar logistik yang semakin kompleks, PT Eka Satya Puspita perlu mengakselerasi transformasi digitalnya, memperluas jaringan distribusi,

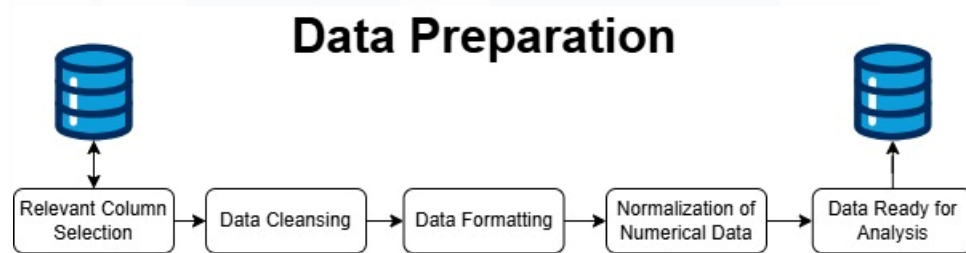
dan meningkatkan layanan monitoring yang selama ini sudah menjadi keunggulan. Dengan pemahaman menyeluruh terhadap lanskap kompetitor, PT Eka Satya Puspita diharapkan mampu memformulasikan strategi pengembangan yang lebih tajam, baik dalam aspek operasional, pemasaran, maupun kolaborasi lintas platform.

3.2.2 Eksplorasi dan Validasi Sumber Data (Data Understanding)

Pada tahap *Data Understanding*, fokus utama adalah melakukan eksplorasi awal terhadap sumber data yang tersedia serta memastikan validitas dan relevansinya untuk mendukung proses analisis selanjutnya. Bagi PT Eka Satya Puspita, pemahaman yang mendalam terhadap sumber data menjadi fondasi penting, mengingat seluruh proses operasional pengiriman barang melibatkan berbagai tahapan dan aktor yang beragam. Sebagai bagian dari tahapan ini, proses *pre-processing* data menjadi elemen kunci untuk menjamin kualitas dan konsistensi data sebelum dilakukan analisis lebih lanjut. Selain itu, pemetaan awal terhadap pola dan anomali dalam data juga dilakukan guna mengidentifikasi potensi masalah atau insight yang dapat memengaruhi strategi analisis ke depan. Langkah ini tidak hanya membantu dalam mengidentifikasi potensi masalah seperti data duplikat, nilai yang hilang, atau inkonsistensi format, tetapi juga memberikan insight awal yang dapat digunakan untuk menentukan pendekatan analisis yang paling sesuai dengan konteks operasional perusahaan.

Proses ini melibatkan serangkaian langkah teknis yang dilaksanakan secara sistematis agar data mentah dapat diubah menjadi format yang siap dianalisis. Langkah-langkah tersebut meliputi ekstraksi data dari berbagai sumber internal, pembersihan data dari inkonsistensi, transformasi format, serta normalisasi data numerik untuk memastikan kesetaraan skala antar fitur. Setiap tahap memiliki peran penting dalam meminimalkan risiko kesalahan analisis dan menjaga integritas data sepanjang proses pengolahan. Selain itu, proses ini juga memperhatikan aspek efisiensi agar

data yang dihasilkan tidak hanya bersih, tetapi juga siap untuk digunakan dalam algoritma pemodelan seperti K-Means. Sebagai gambaran, Gambar 3.4 berikut menyajikan diagram alur keseluruhan *pre-processing* yang diterapkan dalam penelitian ini.



Gambar 3. 4 Diagram Alur *Data-Preparation*

Gambar 3.4 diagram di atas menggambarkan enam tahapan utama dalam proses *data preparation*. Tahapan pertama adalah *Data Extraction*, yaitu pengambilan data mentah dari berbagai sumber internal yang tersedia di PT Eka Satya Puspita. Proses dilanjutkan dengan *Data Cleansing* untuk membersihkan data dari nilai kosong, duplikat, dan inkonsistensi yang dapat memengaruhi akurasi analisis. Setelah itu dilakukan *Data Formatting*, yakni penyesuaian struktur data seperti format tanggal dan tipe data agar seragam dan siap diproses lebih lanjut. Berikutnya adalah *Encoding Categorical Data*, yaitu mengubah data kategorikal seperti kota tujuan dan status pengiriman ke dalam format numerik menggunakan teknik *label encoding* atau *one-hot encoding*. Kemudian dilakukan *Normalization of Numerical Data* agar skala antar fitur numerik menjadi seimbang, sehingga algoritma seperti K-Means dapat bekerja secara optimal. Setelah seluruh tahapan selesai, data berada pada kondisi *Data Ready for Analysis* dan siap digunakan untuk proses modeling dan analisis selanjutnya.

1. Mengumpulkan data dari berbagai sumber *internal* dan eksternal

Proses ini diawali dengan identifikasi terhadap dua kategori besar sumber data, yakni sumber data *internal* dan sumber data eksternal. Sumber data *internal* mengacu pada semua catatan operasional yang dimiliki dan dikelola langsung oleh PT Eka Satya Puspita. Salah satu basis data utama

[illegible]

Sebagaimana ditunjukkan dalam struktur data pada Gambar 3.5, tabel *SLA Performance* ini berfungsi sebagai sumber utama yang mencatat seluruh aktivitas penting dalam proses pengiriman. Setiap kolom dalam tabel ini memiliki peran strategis dalam mendukung pemantauan performa dan pelacakan shipment secara real-time maupun historis. Misalnya, kolom *Date* mencatat waktu setiap proses logistik berlangsung, memberikan garis waktu yang akurat dalam mengevaluasi ketepatan pengiriman. Kolom *Vendor*, *SC Origin*, dan *SC Destination* membantu dalam memetakan rantai distribusi barang dari titik awal hingga ke tujuan akhir, sehingga mempermudah dalam analisis jalur-jalur yang berkontribusi pada keterlambatan maupun percepatan pengiriman. Nomor referensi pengiriman

seperti *LT Number* dan *TO Number* memberikan identitas unik bagi setiap transaksi pengiriman, sedangkan *Gross Weight* memberikan informasi kuantitatif yang menjadi dasar dalam perhitungan biaya kirim maupun pemenuhan standar maskapai. Kolom *AWB (Air Waybill)* menjadi elemen penting yang tidak hanya berfungsi sebagai alat pelacakan, namun juga menjadi dokumen legal yang menghubungkan proses pengiriman udara dengan database *internal* perusahaan.

Tabel ini juga mendokumentasikan keterlibatan pihak maskapai secara detail melalui kolom *Actual Airline Name* dan *Flight No*, yang dapat digunakan untuk mengidentifikasi performa maskapai tertentu dalam mendukung pengiriman barang. Waktu keberangkatan dan kedatangan pesawat yang tercatat di kolom *ATD Flight* dan *ATA Flight* memungkinkan tim untuk mengukur ketepatan jadwal dan mengkaji penyebab potensi keterlambatan. Sementara itu, proses pergudangan juga terekam secara rinci melalui kolom *ATA POD Warehouse*, *ATD POD Warehouse*, dan *ATA Vendor WH Destination*, sehingga perjalanan barang dari pesawat ke gudang dapat dilacak tanpa celah. *Kolom Dooring Activity Vendor* dan *Arrival Status Vendor* menjadi indikator langsung untuk memantau keberhasilan proses distribusi di titik akhir, sementara kolom *Issue* mencatat segala kendala yang muncul, seperti masalah barang suspek *Dangerous Goods (DG)* ataupun kendala administratif. Informasi tambahan di kolom *Remarks (Mandatory)* memberikan konteks penting yang membantu tim analisis dalam melakukan investigasi lebih lanjut terhadap kasus-kasus tertentu.

Dalam hal pengukuran kinerja, perbandingan antara *Target SLA Hour* dan *Actual SLA Hour* menjadi pusat perhatian, karena ini secara langsung merefleksikan seberapa baik proses pengiriman memenuhi standar waktu yang telah disepakati. Ketika terjadi gap antara target dan aktual, analisis mendalam dapat dilakukan dengan mengacu pada atribut lain seperti *Station*, *Concat Origin-Station*, serta data tarif dan biaya di kolom

Rate dan Amount. Sementara itu, dari sisi data eksternal, PT Eka Satya Puspita juga memanfaatkan informasi yang diperoleh dari kolaborasi dengan mitra, terutama maskapai penerbangan. Data eksternal ini tidak hanya penting untuk pelacakan status slot pesawat, tetapi juga menjadi dokumen transparansi yang krusial ketika proses *invoicing* dilakukan antara perusahaan dan pihak maskapai. Dengan adanya data slot reservasi yang tervalidasi, proses rekonsiliasi tagihan bisa berlangsung secara akurat dan adil, menghindari potensi selisih biaya yang kerap terjadi dalam transaksi logistik udara. Seluruh proses pengumpulan data ini menjadi fondasi awal yang kuat, namun langkah selanjutnya yang tidak kalah penting adalah memastikan bahwa data yang dikumpulkan benar-benar relevan dan akurat. Validasi data menjadi tahap krusial di sini, untuk menyaring informasi yang layak dianalisis dan mengeliminasi potensi kesalahan yang dapat mengganggu hasil evaluasi performa pengiriman. Oleh karena itu, proses *Data Understanding* di PT Eka Satya Puspita selalu diakhiri dengan tahapan pengecekan integritas dan kesesuaian data sebelum melangkah lebih jauh ke tahap analisis.

2. Memastikan data yang dikumpulkan relevan dan akurat

Setelah proses pengumpulan data dari berbagai sumber selesai, langkah berikutnya yang tak kalah penting dalam tahap *Data Understanding* adalah memastikan bahwa data yang telah dikumpulkan benar-benar relevan dan akurat. Di lingkungan operasional PT Eka Satya Puspita, validasi data ini menjadi hal yang sangat krusial karena seluruh proses analisis performa pengiriman dan pengambilan keputusan strategis bertumpu pada ketepatan data yang dimiliki. Dalam konteks data *internal*, validasi biasanya dimulai dari pengecekan konsistensi antar kolom dalam tabel *SLA Performance*. Misalnya, antara tanggal pada kolom *Date*, waktu keberangkatan (*ATD Flight*), dan waktu kedatangan (*ATA Flight*), semua harus berjalan secara logis dan berurutan. Tim analis juga melakukan pengecekan silang terhadap nomor referensi seperti *AWB*, *LT Number*, dan

TO Number untuk memastikan bahwa setiap shipment tercatat hanya satu kali tanpa duplikasi data. Keakuratan kolom *Gross Weight* juga dikonfirmasi dengan data yang diterima dari pihak maskapai, memastikan tidak ada perbedaan yang dapat mempengaruhi biaya dan beban pesawat.

Pengecekan validitas data eksternal yang diterima dari mitra seperti maskapai dan vendor juga menjadi bagian penting dari proses ini. Data slot reservasi pesawat yang masuk ke sistem PT Eka Satya Puspita dicek ulang terhadap manifest penerbangan resmi yang digunakan untuk keperluan *invoicing*. Hal ini memastikan bahwa setiap transaksi dan tagihan dapat dipertanggungjawabkan secara transparan, baik untuk *internal* manajemen maupun saat proses audit. Selain aspek keakuratan, relevansi data juga diperhatikan dengan seksama. Tidak semua data yang terkumpul selalu langsung digunakan dalam analisis utama. Oleh karena itu, tim melakukan proses filter untuk menyaring hanya data yang berkaitan langsung dengan indikator performa pengiriman, seperti *Target SLA Hour*, *Actual SLA Hour*, *Arrival Status Vendor*, dan *Issue* yang muncul di lapangan. Data yang bersifat *noise* atau tidak memberikan kontribusi signifikan terhadap tujuan analisis akan dipisahkan agar tidak mengganggu akurasi hasil.

Tahap validasi ini juga mencakup proses pemeriksaan manual terhadap data yang memiliki *Issue* atau remark khusus. Data dengan status mencurigakan seperti keterlambatan ekstrem, adanya barang yang dicurigai sebagai *Dangerous Goods* (DG), atau terjadi mismatch antara tujuan pengiriman dan stasiun transit akan ditandai untuk dianalisis lebih lanjut pada tahap berikutnya. Dengan pendekatan validasi yang ketat ini, PT Eka Satya Puspita memastikan bahwa seluruh data yang masuk ke tahap analisis berikutnya adalah data yang *clean*, akurat, dan relevan dengan tujuan bisnis perusahaan. Setelah proses penyaringan dan validasi ini selesai, maka data yang telah terverifikasi siap untuk diproses lebih lanjut ke dalam tahap berikutnya, yaitu *Data Preparation*, di mana data akan dibersihkan secara teknis dan diolah ke dalam format yang siap dianalisis.

3.2.3 Persiapan Data untuk Analisis Lanjutan (Data Preparation)

Setelah proses pemahaman data dan validasi kelengkapan serta akurasi selesai dilakukan, tahap berikutnya dalam siklus analisis di PT Eka Satya Puspita adalah *Data Preparation*. Pada tahap ini, data yang sudah dipastikan relevan akan diproses lebih lanjut agar benar-benar siap digunakan dalam proses analisis lanjutan maupun pemodelan. Tahap ini menjadi sangat penting karena kualitas hasil analisis sangat dipengaruhi oleh bagaimana data disiapkan baik dari sisi kebersihan, format, maupun strukturnya. Dalam pelaksanaannya, *Data Preparation* di PT Eka Satya Puspita mencakup beberapa aktivitas utama, mulai dari penerapan aplikasi analisis data, proses pembersihan data, hingga transformasi data ke dalam format yang sesuai untuk keperluan analisis lebih dalam. Proses ini memastikan bahwa data shipment yang diolah benar-benar mencerminkan kondisi operasional sebenarnya dan dapat diolah secara optimal menggunakan tools analisis yang ada. Berikut adalah tahapan yang dilakukan:

1. Aplikasi Analisis Data (Excel & Python)

Dalam proses *Data Preparation*, tim analis PT Eka Satya Puspita mengandalkan kombinasi antara aplikasi Microsoft Excel dan bahasa pemrograman Python. Excel digunakan terutama untuk proses awal seperti *quick check*, filter data menggunakan *pivot table*, dan pengecekan visual terhadap anomali data, karena kemudahan penggunaannya dalam menavigasi tabel shipment yang besar. Sementara itu, untuk kebutuhan yang lebih kompleks, terutama dalam hal pemodelan dan *data wrangling*, Python menjadi tools utama yang digunakan. Dengan menggunakan library seperti *pandas*, *numpy*, dan *openpyxl*, proses pembersihan data dalam jumlah besar dapat dilakukan secara lebih efisien. Python juga dimanfaatkan dalam proses segmentasi data shipment, perhitungan metrik performa *SLA*, serta penyusunan data dalam format yang siap diproses untuk visualisasi lebih lanjut. Pendekatan ini memberikan fleksibilitas yang lebih tinggi dan

membantu tim untuk membangun pondasi data yang kuat sebelum masuk ke tahap modeling dan analisis *dashboard* interaktif.

2. Melakukan Data Cleaning

Setelah memastikan bahwa aplikasi analisis yang digunakan sudah tepat, langkah berikutnya dalam tahapan *Data Preparation* adalah proses *Data Cleaning*. Proses ini menjadi krusial karena data shipment yang dimiliki PT Eka Satya Puspita berasal dari berbagai sumber dan tahap operasional, sehingga potensi adanya data yang kurang lengkap, duplikat, atau bahkan anomali cukup tinggi. *Data Cleaning* bertujuan untuk menyaring data mentah agar hanya data yang benar-benar akurat, bersih, dan siap pakai yang dilibatkan dalam proses analisis lebih lanjut. Dengan data yang bersih, hasil analisis dapat memberikan gambaran yang lebih akurat terkait performa *SLA* maupun efektivitas pengiriman di tiap rute yang dijalankan.

Sebelum masuk ke tahap *Data Cleaning* secara menyeluruh, langkah awal yang dilakukan adalah memilih kolom-kolom yang relevan untuk analisis. Pemilihan kolom ini menjadi penting agar proses cleaning fokus pada data yang memiliki pengaruh langsung terhadap performa pengiriman. Pada langkah ini, dipilih beberapa kolom utama yaitu *Gross Weight*, *Target SLA Hour*, *Actual SLA Hour*, *Arrival Status Vendor*, dan *SC Destination*. Kolom-kolom ini dipilih karena mewakili dimensi kunci dalam analisis pengelompokan atau *clustering* yang akan dijalankan, khususnya dalam memahami hubungan antara berat kiriman, waktu *SLA*, status kedatangan, dan tujuan pengiriman. Kelima kolom tersebut tidak hanya relevan secara analitis, tetapi juga memiliki konsistensi data yang relatif baik dari hasil eksplorasi awal. Selain itu, pemilihan kolom ini mempertimbangkan efisiensi pemrosesan data, sehingga proses transformasi dan modeling dapat berjalan lebih optimal. Rincian kode dalam pemilihan kolom utama dapat dilihat pada Gambar 3.6.

```
# Memilih kolom yang relevan untuk clustering
df_selected = df[['Gross Weight', 'Target SLA Hour', 'Actual SLA Hour', 'Arrival Status Vendor', 'SC Destination']]

# Menampilkan beberapa baris pertama dari data yang sudah dipilih
df_selected.head()
```

Gambar 3. 6 Proses Pemilihan Kolom yang Relevan

Hasil dari proses pemilihan kolom pada Gambar 3.6 ditunjukkan pada Tabel 3.3 berikut.

Tabel 3. 3 Hasil Pemilihan Kolom yang Relevan

No	Gross Weight	Target <i>SLA</i> Hour	Actual <i>SLA</i> Hour	Arrival Status Vendor	SC Destination
0	21,12	28.0	23.016667	On Time	Makassar DC
1	20,94	28.0	23.016667	On Time	Makassar DC
2	17,87	28.0	23.016667	On Time	Makassar DC
3	23,17	28.0	23.016667	On Time	Makassar DC
4	16,85	28.0	23.016667	On Time	Makassar DC

Table 3.3 dapat dilihat bahwa lima kolom utama telah diseleksi untuk dilanjutkan ke proses *clustering* selanjutnya. Proses seleksi ini penting untuk memastikan bahwa analisis tidak tercampur dengan variabel yang kurang relevan atau bahkan bisa mengganggu pola data. Sebagai contoh, kolom Gross Weight akan membantu mengidentifikasi apakah ada kecenderungan keterlambatan pengiriman pada barang dengan berat tertentu. Sementara kolom Target *SLA* Hour dan Actual *SLA* Hour menjadi tolok ukur utama dalam mengukur ketepatan waktu pengiriman. Sedangkan *Arrival Status Vendor* menunjukkan performa aktual di lapangan, apakah barang sampai *on time* atau mengalami *delay*, dan kolom *SC Destination* memungkinkan analisis berbasis rute tujuan yang spesifik. Strategi preprocessing yang dilakukan dalam penelitian ini mengacu pada pendekatan optimasi pemilihan fitur seperti yang dijelaskan dalam penelitian “*Feature Selection Using New Version of V-Shaped Transfer Function for Salp Swarm Algorithm in Sentiment Analysis*”, di mana tahap seleksi fitur menjadi krusial untuk meningkatkan efisiensi pemodelan dan interpretasi visual [13]. Langkah pemilihan kolom ini juga sekaligus menjadi pondasi bagi proses *Data Cleaning* berikutnya, di mana setiap

kolom akan diperiksa lebih dalam untuk memastikan tidak ada nilai kosong (*missing value*), duplikasi data, atau anomali yang bisa mempengaruhi hasil analisis. Proses pembersihan ini diawali dengan penanganan nilai null yang ditemukan dalam dataset, seperti yang ditunjukkan pada Gambar 3.7.

```
# Mengecek apakah ada nilai null (NaN) di DataFrame
null_check = df_selected.isnull().sum()

# Menampilkan jumlah nilai null per kolom
print(null_check)
```

Gross Weight	0
Target SLA Hour	0
Actual SLA Hour	0
Arrival Status Vendor	0
SC Destination	0
dtype:	int64

Gambar 3. 7 Proses Penanganan Nilai Null

Gambar 3.7 pada tahap *Data Cleaning*, langkah penting yang dilakukan selanjutnya adalah proses menangani nilai *null* (kosong) yang ditemukan dalam dataset. Seperti yang terlihat pada Gambar 3.6 Proses Menangani Nilai *Null*, proses ini diawali dengan perintah penghapusan baris yang mengandung nilai *NaN* (*Not a Number*) menggunakan fungsi *dropna()*. Tahapan ini dilakukan guna memastikan bahwa seluruh informasi yang dipakai dalam proses pengolahan data sudah benar-benar utuh dan tidak terdapat kekosongan yang dapat mempengaruhi kualitas model. Mengingat data yang digunakan bersifat operasional shipment, keberadaan nilai kosong bisa menimbulkan bias dalam perhitungan *SLA* maupun performa rute pengiriman. Setelah proses penghapusan, dicek kembali seluruh kolom menggunakan fungsi *isnull ()*. *sum ()* untuk memastikan bahwa tidak ada lagi kolom yang mengandung nilai kosong.

Hasil pengecekan menunjukkan bahwa semua kolom utama mulai dari Gross Weight, Target *SLA* Hour, Actual *SLA* Hour, Arrival Status

Vendor, hingga SC Destination telah bersih, masing-masing menunjukkan angka nol pada jumlah nilai null-nya. Ini menandakan bahwa dataset sudah lolos tahap validasi data kosong dan siap masuk ke tahap pengolahan berikutnya. Dengan data yang kini sudah bersih dari nilai kosong, proses *Data Preparation* dapat dilanjutkan ke tahap berikutnya, yaitu mengolah data ke format yang siap dianalisis. Tahap ini bertujuan untuk memastikan bahwa data yang telah dibersihkan tidak hanya lengkap, tetapi juga berada dalam format numerik atau kategorikal yang sesuai untuk dimasukkan ke dalam model analitik atau algoritma *clustering* yang akan digunakan.

3. Mengolah Data ke Format yang Siap Dianalisis

Setelah proses *data cleaning* selesai dan dataset sudah bebas dari nilai kosong maupun data duplikat, tahap selanjutnya dalam *data preparation* adalah mengolah data ke dalam format yang siap dianalisis menggunakan metode *clustering K-Means*. Salah satu karakteristik penting dari algoritma *K-Means* adalah kemampuannya yang hanya terbatas pada pemrosesan data numerik. Oleh karena itu, sebelum memasuki tahap pemodelan, diperlukan proses transformasi terhadap kolom-kolom yang masih bertipe kategorikal agar dapat dibaca dan diproses oleh algoritma. Pada tahap ini, perhatian utama diarahkan pada kolom *SC Destination*, yang berisi nama-nama kota tujuan dari proses distribusi barang. Kolom ini termasuk jenis data kategorikal yang tidak dapat diproses langsung tanpa dikonversi ke dalam format angka.

Transformasi dilakukan menggunakan teknik *One-Hot Encoding*, yaitu metode yang mengubah setiap kategori unik menjadi vektor biner tersendiri, sehingga hubungan antar kategori tidak disalahartikan sebagai nilai ordinal. Proses ini menghasilkan beberapa kolom baru yang masing-masing mewakili satu kota tujuan, dengan nilai 1 jika baris data terkait mengarah ke kota tersebut, dan 0 jika tidak. Proses *One-Hot Encoding* ini divisualisasikan pada Gambar 3.8, yang menunjukkan struktur data setelah transformasi dilakukan. Dengan demikian, data hasil *encoding* ini telah

sesuai dengan kebutuhan input model *K-Means* dan siap untuk dianalisis lebih lanjut dalam tahap pemodelan kluster.

```
# One-Hot Encoding untuk SC Destination
df_encoded = pd.get_dummies(df_selected, columns=['SC Destination'])

# Menampilkan beberapa baris setelah encoding
df_encoded.head()
```

	Gross Weight	Target SLA Hour	Actual SLA Hour	Arrival Status Vendor	SC Destination_Aceh DC	SC Destination_Alak DC	SC Destination_Baguala DC	SC Destination_Balikpapan DC	SC Destination_Batam DC	SC Destination_Darul Imarah DC	...	Desti
0	21.12	28.0	23.016667	On Time	False	False	False	False	False	False	...	
1	20.94	28.0	23.016667	On Time	False	False	False	False	False	False	...	
2	17.87	28.0	23.016667	On Time	False	False	False	False	False	False	...	
3	23.17	28.0	23.016667	On Time	False	False	False	False	False	False	...	
4	16.85	28.0	23.016667	On Time	False	False	False	False	False	False	...	

5 rows × 27 columns

Gambar 3. 8 Proses One-Hot Encoding untuk Kolom Kategorikal "SC Destination"

Gambar 3.8 *one-hot encoding* untuk kolom kategorikal "*SC Destination*", proses ini mengubah setiap kategori unik dari *SC Destination* menjadi kolom baru yang bersifat biner (*True/False*). Sebagai contoh, jika sebuah baris data memiliki tujuan ke *Makassar DC*, maka kolom *SC Destination_Makassar DC* akan bernilai *True*, sementara kolom tujuan lainnya akan bernilai *False*. Dengan cara ini, informasi lokasi tujuan tetap dapat dipertahankan dalam dataset tanpa melanggar aturan numerik yang diperlukan oleh *K-Means*. Hasil encoding ini menambah banyak kolom baru ke dalam dataset, masing-masing mewakili satu kota atau hub distribusi di jaringan ESP, seperti *Aceh DC*, *Pontianak DC*, *Ternate Hub*, dan lainnya. Langkah ini sangat penting untuk memastikan bahwa faktor tujuan pengiriman ikut diperhitungkan dalam pengelompokan rute, namun dalam bentuk data yang kompatibel dengan algoritma. Setelah transformasi *One-Hot Encoding* selesai diterapkan untuk kolom tujuan, selanjutnya kita juga perlu menangani kolom kategorikal lainnya, yaitu *Arrival Status Vendor*. Pada tahap berikutnya, kolom ini akan kita ubah menggunakan teknik *Label Encoding*, seperti yang akan diperlihatkan dalam Gambar 3.9 *Label Encoding* untuk *Arrival Status Vendor*.

```

from sklearn.preprocessing import LabelEncoder

# Label Encoding untuk Arrival Status Vendor
label_encoder = LabelEncoder()
df_encoded['Arrival Status Vendor'] = label_encoder.fit_transform(df_encoded['Arrival Status Vendor'])

# Menampilkan beberapa baris setelah encoding
df_encoded.head()

```

	Gross Weight	Target SLA Hour	Actual SLA Hour	Arrival Status Vendor	Destination_Aceh DC	SC Destination_Alak DC	SC Destination_Bagualla DC	SC Destination_Balikpapan DC	SC Destination_Batam DC	SC Destination_Darul Imarah DC	...	Desti
0	21.12	28.0	23.016667	1	False	False	False	False	False	False	...	...
1	20.94	28.0	23.016667	1	False	False	False	False	False	False	...	...
2	17.87	28.0	23.016667	1	False	False	False	False	False	False	...	...
3	23.17	28.0	23.016667	1	False	False	False	False	False	False	...	...
4	16.85	28.0	23.016667	1	False	False	False	False	False	False	...	...

5 rows x 27 columns

Gambar 3. 9 Proses *Label Encoding* untuk *Arrival Status Vendor*

Gambar 3.9 setelah kolom *SC Destination* berhasil dikonversi menggunakan teknik *One-Hot Encoding*, langkah selanjutnya adalah menangani kolom kategorikal lainnya yaitu *Arrival Status Vendor*. Kolom ini sebelumnya berisi label berupa teks seperti "*On Time*" dan "*Late*", yang secara langsung tidak dapat digunakan dalam proses perhitungan algoritma *K-Means*. Oleh karena itu, pada tahap ini kita menggunakan teknik *Label Encoding* untuk mengubah nilai-nilai teks tersebut menjadi angka. Seperti yang ditampilkan pada *Gambar 3.8 Label Encoding untuk Arrival Status Vendor*, proses ini mengonversi status pengiriman "*On Time*" menjadi 1, sementara "*Late*" dikodekan menjadi 0. Konversi ini memungkinkan algoritma *K-Means* untuk memahami status pengiriman dalam bentuk numerik yang dapat dihitung dalam proses pengelompokan. Hasilnya, dataset kita kini semakin bersih dan siap untuk masuk ke tahap selanjutnya karena semua kolom kategorikal telah dikonversi menjadi bentuk numerik. Transformasi ini penting untuk memastikan bahwa setiap fitur dalam dataset dapat diproses oleh algoritma *K-Means*, yang mengandalkan perhitungan jarak numerik antar data. Namun, meskipun seluruh kolom kini sudah dalam format numerik, kita masih perlu melakukan satu tahap penting sebelum proses *clustering*, yaitu normalisasi data numerik. Ini bertujuan untuk memastikan bahwa semua kolom memiliki skala yang seragam, sehingga fitur dengan nilai besar seperti *Gross Weight* tidak mendominasi proses pengelompokan. Proses ini akan dijelaskan lebih lanjut dalam Gambar 3.10 yang akan menunjukkan bagaimana nilai-nilai dari kolom numerik

disesuaikan agar setara dalam skala yang sama sebelum masuk ke tahap *K-Means Clustering*.



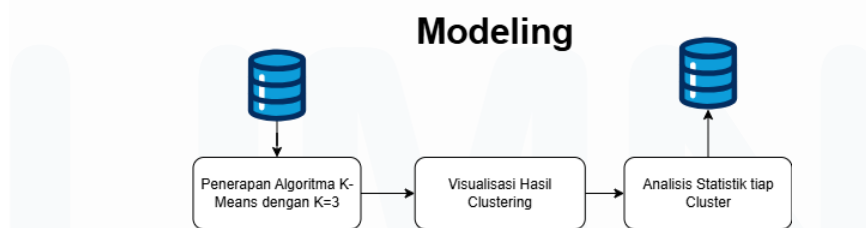
Gambar 3. 10 Proses Normalisasi Data Numerik

Gambar 3.10 tahap terakhir dalam proses *data preparation* adalah normalisasi data numerik. Proses ini penting karena algoritma *clustering* seperti *K-Means* bekerja dengan mengukur jarak antar titik data. Jika skala antar kolom berbeda jauh, fitur dengan rentang angka besar, seperti *Gross Weight*, bisa secara tidak proporsional mempengaruhi hasil pengelompokan. Untuk itu, normalisasi memastikan setiap fitur numerik berkontribusi secara adil dalam proses ini. Dalam langkah ini, kolom numerik yang dipilih untuk normalisasi adalah *Gross Weight*, *Target SLA Hour*, dan *Actual SLA Hour*. Menggunakan metode *Standard Scaler* dari library *scikit-learn*, yang mengubah nilai-nilai tersebut menjadi distribusi dengan rata-rata nol dan standar deviasi satu. Hasilnya bisa dilihat pada tabel output, di mana nilai *Gross Weight* dan *SLA Hour* kini telah berubah menjadi angka yang lebih kecil, sebagian besar berada di kisaran -1 hingga 1. Sebagai contoh, nilai *Gross Weight* pada baris pertama kini bernilai 0.741580, yang berarti sekitar 0.7 standar deviasi di atas rata-rata dari seluruh data pengiriman. Sedangkan nilai *Target SLA Hour* dan *Actual SLA Hour* keduanya kini terstandarisasi sehingga siap dianalisis lebih lanjut tanpa ada dominasi dari skala aslinya. Dengan selesainya proses normalisasi ini, seluruh dataset kini telah sepenuhnya dalam format numerik yang

bersih, relevan, dan siap untuk digunakan dalam proses modeling. Baik kolom kategorikal maupun numerik kini telah diproses dengan baik sehingga bisa langsung diolah menggunakan algoritma *K-Means* untuk mengelompokkan rute pengiriman berdasarkan performa *SLA* dan karakteristik lainnya. Pada tahap berikutnya, yaitu penerapan teknik analisis (Modeling), kita akan mulai menerapkan teknik *clustering* menggunakan *K-Means* untuk mendapatkan pola pengelompokan yang bisa memberikan *insight* baru.

3.2.4 Penerapan Teknik Analisis (Modeling)

Setelah seluruh data dipersiapkan melalui proses encoding dan normalisasi, tahap selanjutnya adalah penerapan algoritma K-Means untuk mengelompokkan data dan menemukan pola tersembunyi dalam performa pengiriman. Proses ini mencakup tiga langkah utama yaitu klasterisasi data, visualisasi hasil clustering, serta analisis statistik tiap cluster untuk memahami karakteristik pengiriman. Ketiga tahapan tersebut disusun secara sistematis guna mengidentifikasi pola berdasarkan kemiripan atribut, seperti divisualisasikan pada Gambar 3.11.



Gambar 3. 11 Diagram Alur Proses *Modeling*

3.2.4.1 Penerapan Algoritma *K-Means* dengan $K=3$

Dalam penelitian ini, teknik yang digunakan adalah *K-Means Clustering*, sebuah metode *unsupervised learning* yang efektif untuk membagi data ke dalam kelompok (*cluster*) berdasarkan kesamaan karakteristik. Tujuan utama dari penerapan *K-Means* ini adalah untuk mengelompokkan rute pengiriman dan vendor berdasarkan performa *SLA*, berat kiriman, dan status kedatangan. Dengan pengelompokan ini, kita dapat

mengidentifikasi mana saja rute yang cenderung bermasalah (*late delivery*), mana yang berjalan normal, dan mana yang memiliki karakteristik unik lainnya. Hasil dari *clustering* ini juga diharapkan akan dijadikan insight operasional, sehingga tim operasional PT ESP dapat memantau kondisi pengiriman secara berkala dan mengambil tindakan lebih cepat untuk rute yang masuk dalam *cluster* berisiko. Pada tahap awal modeling, kita menerapkan *K-Means* dengan jumlah *cluster* sebanyak 3, yang ditampilkan pada Gambar 3.12. Pemilihan angka 3 ini didasarkan pada asumsi awal bahwa performa pengiriman bisa dibagi menjadi tiga kategori besar: baik (*on-time*), sedang (*borderline*), dan buruk (*late delivery*).

```
from sklearn.cluster import KMeans

# Melakukan K-Means Clustering dengan 3 klaster
kmeans = KMeans(n_clusters=3, random_state=42)
df_encoded['Cluster'] = kmeans.fit_predict(df_encoded)

# Menampilkan hasil clustering
print(df_encoded[['Gross Weight', 'Target SLA Hour', 'Actual SLA Hour', 'Arrival Status Vendor', 'Cluster']].head())
```

	Gross Weight	Target SLA Hour	Actual SLA Hour	Arrival Status Vendor	\
0	0.741580	-0.599187	-0.068452	1	
1	0.720303	-0.599187	-0.068452	1	
2	0.357409	-0.599187	-0.068452	1	
3	0.983904	-0.599187	-0.068452	1	
4	0.236838	-0.599187	-0.068452	1	

	Cluster
0	2
1	2
2	2
3	2
4	2

Gambar 3.12 Proses *K-Means Clustering* dengan $K=3$

Gambar 3.12 berdasarkan hasil output yang menunjukkan bahwa setiap pengiriman kini sudah memiliki label *Cluster*. Sebagai contoh, pada lima baris pertama, seluruh pengiriman mendapatkan label *cluster 2*, yang berarti menurut algoritma *K-Means*, data tersebut masuk ke kelompok yang memiliki karakteristik serupa. Jika kita perhatikan lebih dalam, kolom *Gross Weight* sudah dalam bentuk ternormalisasi dengan variasi yang cukup terlihat, menandakan adanya perbedaan beban kiriman antar shipment. Sementara itu, kolom *Target SLA Hour* yang juga sudah dinormalisasi, nilainya beragam karena memang *Target SLA* ditentukan berbeda-beda tergantung pada destinasi masing-masing shipment. Misalnya, destinasi yang jauh seperti Indonesia Timur biasanya memiliki target *SLA* yang lebih

tinggi dibandingkan destinasi yang lebih dekat. Hal ini menciptakan variasi target *SLA* yang menjadi salah satu penentu penting dalam proses *clustering* ini.

Kolom *Actual SLA Hour* mencerminkan perbedaan waktu aktual pengiriman antar *shipment*, yang dibandingkan dengan target *SLA* untuk menilai performa distribusi. Selisih ini memberikan indikasi apakah pengiriman dilakukan tepat waktu atau mengalami keterlambatan, sehingga sangat relevan untuk menilai efektivitas layanan logistik secara menyeluruh. Sementara itu, kolom *Arrival Status Vendor* yang telah melalui proses *Label Encoding* menghasilkan nilai 1 untuk status *On Time*, menandakan bahwa pengiriman tersebut berhasil memenuhi target *SLA* yang telah ditentukan.

Manfaat penggunaan teknik K-Means ini, pengelompokan *shipment* menjadi lebih terstruktur, kita bisa mengamati apakah suatu cluster cenderung didominasi oleh destinasi dengan target *SLA* tinggi, atau mungkin cluster lain diisi oleh *shipment* yang sering terlambat (*late delivery*) meskipun target *SLA*-nya rendah. Pola-pola ini nantinya bisa dijadikan landasan dalam menyusun rekomendasi peningkatan layanan logistik yang lebih spesifik per rute atau destinasi. Sebagai langkah selanjutnya, hasil *clustering* ini akan divisualisasikan untuk memperoleh gambaran yang lebih jelas mengenai distribusi dan karakteristik tiap cluster, yang akan dibahas pada subbab berikutnya, yaitu 3.2.4.2 Visualisasi Hasil *Clustering*.

3.2.4.2 Proses Visualisasi Hasil Clustering

Tahap selanjutnya setelah penerapan algoritma K-Means adalah visualisasi hasil *clustering* untuk memahami bagaimana data pengiriman terdistribusi ke dalam masing-masing cluster. Visualisasi ini bertujuan memberikan gambaran awal terhadap pola yang terbentuk berdasarkan

atribut operasional yang relevan. Proses ini divisualisasikan secara grafis seperti ditunjukkan pada Gambar 3.13.



Gambar 3. 13 Proses Visualisasi Hasil *Clustering*

Visualisasi hasil *clustering* yang ditampilkan pada Gambar 3.13 memberikan gambaran yang cukup jelas tentang bagaimana data pengiriman dikelompokkan berdasarkan kinerja operasionalnya. Dalam grafik ini, setiap titik mewakili satu data pengiriman, yang diplot berdasarkan nilai *Gross Weight* (berat kotor) pada sumbu X dan *Actual SLA Hour* (jam realisasi *SLA*) pada sumbu Y. Dari hasil visualisasi ini, terlihat bahwa terdapat tiga *cluster* yang terbentuk, masing-masing diberi warna yang berbeda. *Cluster 0* yang ditampilkan dengan warna ungu tampak mendominasi wilayah bawah grafik, yaitu area dengan nilai *Actual SLA Hour* yang rendah. Hal ini menunjukkan bahwa kelompok ini cenderung berisi data pengiriman yang tepat waktu atau bahkan lebih cepat dari *SLA* yang ditargetkan. Klaster ini tampaknya menjadi kelompok dengan performa terbaik dalam hal pemenuhan *SLA*. Selanjutnya, *Cluster 1* yang

berwarna hijau kebiruan juga mengisi area bawah grafik, namun jika diperhatikan lebih seksama, titik-titiknya cenderung terkonsentrasi di area dengan *Gross Weight* yang lebih rendah. Pola ini mengindikasikan bahwa shipment dengan bobot ringan membentuk kelompok tersendiri yang memiliki kecenderungan *SLA* aktual yang baik, namun dengan karakteristik operasional yang berbeda dari shipment dengan bobot lebih besar. Menariknya, perbedaan antara *Cluster 0* dan *Cluster 1* ini bisa saja mencerminkan variasi proses logistik atau jenis layanan yang digunakan tergantung pada berat muatan.

Cluster 2 yang divisualisasikan dengan warna kuning tampak sangat menonjol, karena sebagian besar titiknya tersebar di area atas grafik, tepatnya pada nilai *Actual SLA Hour* yang tinggi, berkisar antara 12.5 hingga 17.5 jam. Pola ini dengan jelas menandakan adanya kelompok shipment yang sering kali mengalami keterlambatan atau realisasi *SLA* yang jauh di atas target. Kondisi ini dapat menjadi perhatian khusus bagi pihak manajemen logistik, karena *cluster* ini merepresentasikan potensi *pain point* dalam performa pengiriman, yang perlu diidentifikasi penyebabnya lebih lanjut apakah disebabkan oleh rute tertentu, kondisi geografis, beban muatan yang berat, atau faktor lain di luar kendali. Secara umum, visualisasi ini sudah memberikan sinyal awal tentang bagaimana performa pengiriman terbagi ke dalam kelompok-kelompok dengan karakteristik yang berbeda. Namun, untuk dapat memahami lebih dalam mengenai apa yang sebenarnya membedakan setiap *cluster* ini, tentu kita memerlukan analisis lanjutan yang lebih berbasis angka. Oleh karena itu, langkah berikutnya adalah memeriksa statistik tiap *cluster* secara lebih rinci, dengan melihat rata-rata dan variasi pada variabel-variabel penting seperti *Gross Weight*, *Target SLA Hour*, *Actual SLA Hour*, hingga *Arrival Status Vendor* pada masing-masing kelompok. Dengan pendekatan ini, kita bisa memperoleh gambaran yang lebih konkret dan komprehensif terkait performa serta tantangan dalam tiap segmentasi data hasil clustering, yang akan dibahas pada subbab 3.2.4.3 Analisis Statistik Setiap Cluster.

3.2.4.3 Analisis Statistik Setiap Cluster.

Tahapan selanjutnya yaitu analisis statistik dilakukan untuk memahami karakteristik tiap cluster secara lebih mendalam setelah tahap visualisasi. Pemeriksaan ini mencakup nilai rata-rata dan standar deviasi dari variabel penting seperti Gross Weight, Target SLA Hour, Actual SLA Hour, dan Arrival Status Vendor. Hasil rekapitulasi statistik tersebut ditampilkan pada Gambar 3.14.

```
# Memeriksa statistik setiap klaster
print(df_encoded.groupby('Cluster').agg({
    'Gross Weight': ['mean', 'std'],
    'Target SLA Hour': ['mean', 'std'],
    'Actual SLA Hour': ['mean', 'std'],
    'Arrival Status Vendor': ['mean', 'std']
}))
```

	Gross Weight		Target SLA Hour		Actual SLA Hour \
	mean	std	mean	std	mean
Cluster					
0	0.713270	0.581517	0.785660	0.874204	-0.068427
1	-1.056583	0.395933	0.008885	0.732237	-0.068408
2	0.594470	0.676887	-0.886416	0.631792	0.166090

	Arrival Status Vendor		
	std	mean	std
Cluster			
0	0.000164	0.941216	0.235229
1	0.000171	0.861852	0.345068
2	1.840924	0.725398	0.446335

Gambar 3. 14 Proses Memeriksa Hasil Klaster Secara Menyeluruh

Gambar 3.14 setelah visualisasi memberikan gambaran umum terkait pemisahan *cluster*, langkah berikutnya adalah melakukan pemeriksaan statistik secara lebih mendalam terhadap tiap *cluster* yang telah terbentuk. Hasil dari analisis statistik ini menunjukkan adanya perbedaan karakteristik yang cukup signifikan antar ketiga *cluster*. Cluster 0, misalnya, memiliki rata-rata *Gross Weight* sebesar 0.71 (dalam nilai yang telah dinormalisasi), dengan deviasi standar sebesar 0.58. Nilai rata-rata *Target SLA Hour* pada cluster ini mencapai 0.78, yang menunjukkan bahwa pengiriman dalam kelompok ini umumnya memiliki target *SLA* yang cukup tinggi dibanding *cluster* lainnya. Namun menariknya, *Actual SLA Hour*-nya

rata-rata berada di angka -0.06, yang berarti realisasi pengiriman cenderung lebih cepat atau tepat waktu, sesuai atau bahkan lebih baik dari target. Hal ini diperkuat oleh nilai rata-rata *Arrival Status Vendor* sebesar 0.94, yang menunjukkan bahwa hampir seluruh pengiriman dalam *cluster* ini dikategorikan sebagai *On Time*. Dapat dikatakan bahwa *Cluster 0* merupakan kelompok dengan performa terbaik dan stabil, dilihat dari rendahnya standar deviasi pada *Arrival Status Vendor* yang hanya sebesar 0.23, mencerminkan konsistensi yang tinggi dalam pencapaian *SLA*.

Cluster 1 memperlihatkan karakteristik yang cukup unik. Rata-rata *Gross Weight* pada *cluster* ini tercatat paling rendah, yaitu -1.05, menunjukkan bahwa kelompok ini didominasi oleh shipment dengan bobot yang ringan. *Target SLA Hour*-nya juga relatif rendah, dengan nilai rata-rata 0.008, yang berarti bahwa pengiriman dalam *cluster* ini biasanya ditargetkan selesai dalam waktu singkat. *Actual SLA Hour*-nya pun rata-rata berada di angka -0.06, serupa dengan *Cluster 0*, yang mengindikasikan bahwa performa realisasi *SLA*-nya tetap baik meskipun bobotnya ringan. Namun, jika ditelusuri lebih lanjut, nilai rata-rata *Arrival Status Vendor* di *cluster* ini sebesar 0.86, sedikit lebih rendah dibanding *Cluster 0*. Ini menunjukkan bahwa meskipun sebagian besar shipment dalam kelompok ini *On Time*, terdapat sedikit ketidakpastian atau risiko keterlambatan yang lebih tinggi ketimbang *cluster* sebelumnya. Hal ini tercermin dari nilai standar deviasi *Arrival Status Vendor* yang juga sedikit lebih besar, yaitu 0.34.

Cluster 2 adalah yang paling menonjol, dari sisi statistik tampak memiliki profil performa yang berbeda drastis. Rata-rata *Gross Weight*-nya tercatat sebesar 0.59, mirip dengan *Cluster 0*, yang berarti kelompok ini mengangkut shipment dengan bobot sedang hingga berat. Namun, *Target SLA Hour*-nya justru bernilai negatif (-0.88), yang mencerminkan bahwa kelompok ini didominasi oleh shipment dengan target *SLA* yang rendah, atau dengan kata lain target penyelesaian yang cepat. Akan tetapi, realisasi

Actual SLA Hour-nya justru jauh lebih tinggi dibanding *cluster* lainnya, dengan rata-rata 0.16. Angka ini mengindikasikan bahwa pengiriman dalam *cluster* ini sering mengalami keterlambatan, dan cenderung gagal memenuhi *SLA* yang ditargetkan. Fakta ini juga tercermin jelas dalam rata-rata *Arrival Status Vendor* yang hanya mencapai 0.72, angka yang jauh lebih rendah dibanding Cluster 0 dan 1. Standar deviasi yang besar pada variabel ini, yaitu 0.44, semakin menegaskan bahwa cluster ini tidak hanya memiliki performa yang buruk, tetapi juga sangat tidak konsisten dalam pencapaian *SLA*-nya.

Dari hasil statistik ini, terlihat bahwa ketiga *cluster* memiliki karakteristik yang unik *Cluster* 0 sebagai kelompok dengan performa *SLA* yang stabil dan baik, *Cluster* 1 sebagai shipment berbobot ringan dengan performa *SLA* yang umumnya baik namun sedikit kurang stabil, dan *Cluster* 2 yang justru menunjukkan performa yang lemah serta ketidakstabilan yang signifikan dalam memenuhi *SLA*. Menariknya, analisis lebih dalam mengungkap bahwa beberapa destinasi pengiriman ternyata masuk ke dalam cluster berisiko (*Cluster* 2), sehingga dapat menjadi perhatian khusus dalam upaya peningkatan performa ke depannya. Informasi lebih rinci mengenai distribusi destinasi dalam masing-masing cluster disajikan pada Tabel 3.4 Rincian Cluster Dominan.

Tabel 3. 4 Rincian Cluster Dominan

No	Destination	Cluster 0	Cluster 1	Cluster 2	Cluster Dominan
1.	SC Destination_Alak DC	29	11	0	Cluster 0
2.	SC Destination_Mataram DC	1791	1432	0	Cluster 0
3.	SC Destination_Padang DC	5375	2708	0	Cluster 0
4	SC Destination_Pekanbaru DC	2230	1117	0	Cluster 0
5.	SC Destination_Tanjung Redeb DC	102	0	0	Cluster 0
6.	SC Destination_Ternate Hub	159	24	0	Cluster 0

No	Destination	Cluster 0	Cluster 1	Cluster 2	Cluster Dominan
7.	SC Destination_Ternate Selatan Hub	177	33	0	Cluster 0
8.	SC Destination_Ternate Utara Hub	106	19	0	Cluster 0
9.	SC Destination_Aceh DC	313	332	0	Cluster 1
10.	SC Destination_Batam DC	0	956	702	Cluster 1
11.	SC Destination_Darul Imarah DC	186	219	0	Cluster 1
12.	SC Destination_Denpasar DC	1355	1726	0	Cluster 1
13.	SC Destination_Makassar DC	0	277	245	Cluster 1
14.	SC Destination_Tamalanrea DC	0	629	613	Cluster 1
15.	SC Destination_Baguala DC	0	33	72	Cluster 2
16.	SC Destination_Balikpapan DC	0	2783	4439	Cluster 2
17.	SC Destination_Dungingi DC	0	312	478	Cluster 2
18.	SC Destination_Kalawat DC	0	408	899	Cluster 2
19.	SC Destination_Palangka Raya DC	0	363	527	Cluster 2
20.	SC Destination_Pangkal Pinang DC	0	26	229	Cluster 2
21.	SC Destination_Pontianak DC	0	208	1590	Cluster 2
22.	SC Destination_Tanjung Pandan DC	0	40	129	Cluster 2
23.	SC Destination_Wua-Wua DC	0	300	696	Cluster 2

Table 3.4 temuan ini menjadi bahan penting untuk memahami risiko operasional secara lebih menyeluruh, meskipun tidak secara langsung divisualisasikan dalam dashboard interaktif. Selain sebagai insight, hasil ini juga dapat dijadikan pertimbangan awal dalam menentukan prioritas rute yang memerlukan evaluasi atau perbaikan. Selanjutnya, penerapan teknik visualisasi dalam bentuk dashboard interaktif akan difokuskan untuk memberikan gambaran performa *SLA* secara komprehensif per destinasi dan secara keseluruhan, agar pihak manajemen dapat dengan mudah memantau,

mengevaluasi, dan mengambil langkah-langkah perbaikan yang lebih terarah. Dengan pendekatan ini, dashboard yang dibangun tetap fokus pada *SLA Performance* yang konkret, sementara hasil *clustering* ini berperan sebagai insight tambahan di belakang layar untuk mendukung pengambilan keputusan yang lebih relevan.

3.2.5 Penilaian Kinerja Model (Evaluation)

Pada tahap ini, dilakukan proses evaluasi terhadap kinerja model *clustering* yang telah dibangun sebelumnya, dengan tujuan untuk mengetahui sejauh mana kualitas segmentasi data yang telah dihasilkan. Evaluasi ini penting untuk memastikan bahwa pembagian data ke dalam masing-masing *cluster* benar-benar mencerminkan perbedaan yang signifikan antar kelompok dan juga memiliki kekompakan yang baik di dalam setiap *cluster*. Salah satu metrik yang digunakan dalam penelitian ini untuk menilai kualitas pembentukan *cluster* adalah Silhouette Score, yang secara umum mengukur seberapa dekat suatu data dengan anggota *cluster* yang sama dibandingkan dengan data di *cluster* lain. Semakin tinggi nilai Silhouette Score, maka semakin baik kualitas pemisahan antar *cluster* dan kekompakan di dalam *cluster*. Dalam konteks penelitian ini, evaluasi dilakukan dengan menggunakan subset data karena keterbatasan memori saat memproses keseluruhan dataset. Meski begitu, pemilihan sampel dilakukan secara cermat agar tetap merepresentasikan variasi data yang ada. Gambar 3.15 menunjukkan hasil evaluasi clustering dengan Silhouette Score terhadap subset data sampel.

```
# Sampling 1000 data dari df_encoded untuk menghitung Silhouette Score
subset_df = df_encoded.sample(n=1000, random_state=42)

# Menghitung Silhouette Score pada subset data
silhouette_avg = silhouette_score(subset_df, subset_df['Cluster'])

# Menampilkan hasil
print(f"Silhouette Score untuk subset data: {silhouette_avg}")

Silhouette Score untuk subset data: 0.35395073649008857
```

Gambar 3. 15 Evaluasi Cluster Menggunakan Silhouette Score

Gambar 3.15 nilai *Silhouette Score* sebesar 0.3539 yang diperoleh dari hasil evaluasi ini menunjukkan bahwa pembentukan cluster yang telah dilakukan berada dalam kategori cukup baik atau *fair* dalam hal pemisahan antar *cluster* dan kekompakan data di dalam setiap *cluster*. Meskipun angka ini belum mencapai kategori tinggi, di mana umumnya nilai di atas 0.5 baru dianggap sangat baik, hasil ini masih dapat diterima dalam konteks penelitian ini. Hal ini disebabkan karena penentuan jumlah *cluster* sebanyak tiga telah ditetapkan sejak awal dengan pendekatan yang berbasis pada kebutuhan bisnis, yaitu untuk mengelompokkan pengiriman barang menjadi tiga tingkatan risiko yang berbeda, mulai dari pengiriman yang aman, pengiriman dengan potensi risiko sedang, hingga pengiriman yang tergolong lebih berbahaya atau memiliki risiko tinggi. Dalam proses penghitungan *Silhouette Score* ini, dilakukan pengambilan sampel sebanyak 1000 data dari total keseluruhan data. Penggunaan sampling ini dilakukan untuk mengatasi keterbatasan memori yang terjadi saat mencoba menghitung *Silhouette Score* pada dataset penuh. Secara teori, penggunaan seluruh data tanpa proses sampling sebenarnya berpotensi menghasilkan nilai *Silhouette Score* yang sedikit lebih tinggi, karena pola distribusi data yang lebih utuh dapat memberikan representasi yang lebih akurat terhadap kualitas pembentukan *cluster*. Namun demikian, hasil penghitungan dengan metode sampling ini sudah cukup untuk memberikan gambaran yang representatif tentang kualitas segmentasi *cluster* yang terbentuk dalam penelitian ini. Perlu juga dipahami bahwa nilai *Silhouette Score* yang tidak terlalu tinggi seperti ini tidak selalu berarti bahwa model clustering yang digunakan gagal atau buruk. Dalam konteks penelitian ini, yang lebih menekankan pada segmentasi risiko pengiriman logistik, tujuan utama bukan semata-mata untuk mencapai pemisahan *cluster* yang maksimal secara statistik, melainkan untuk memetakan risiko pengiriman ke dalam tiga kategori utama sesuai dengan kebutuhan bisnis yang ada. Oleh karena itu, nilai 0.3539 yang diperoleh tetap relevan dan valid untuk

menggambarkan hasil pemetaan risiko pengiriman yang telah dilakukan melalui pendekatan *clustering* ini.

3.2.6 Pembuatan & Implementasi *Dashboard* Interaktif (Deployment)

Setelah proses analisis data dan clustering dilakukan, tahap berikutnya dalam penelitian ini adalah membangun sebuah *dashboard* interaktif yang dapat membantu tim operasional dan manajemen dalam memantau performa pengiriman secara lebih praktis dan berkala. *Dashboard* ini dirancang agar mudah digunakan oleh pengguna non-teknis, dengan fokus utama pada penyajian informasi yang bersifat eksekutif seperti performa *SLA*, analisa keterlambatan, serta *overview* pendapatan perusahaan dari aktivitas pengiriman. Pembuatan dashboard dalam penelitian ini mengacu pada praktik perancangan antarmuka berbasis repository, di mana visualisasi dimanfaatkan untuk menyajikan hasil keluaran model secara efisien melalui sistem berbasis web yang mudah digunakan oleh pengguna [14]. Dalam proses perancangannya, *dashboard* ini tidak secara langsung menampilkan hasil modeling seperti *K-Means Clustering* yang telah dibahas sebelumnya. Hal ini disesuaikan dengan kebutuhan pengguna *dashboard*, yang lebih memerlukan informasi kinerja nyata (misalnya, tingkat pencapaian *SLA* tiap destinasi dan tren keterlambatan) dibandingkan dengan *insight* statistik seperti hasil segmentasi *clustering*. Oleh karena itu, *dashboard* ini lebih diarahkan sebagai alat monitoring performa yang operasional, sementara hasil *clustering* tetap menjadi *background insight* yang dapat digunakan secara internal untuk evaluasi mendalam. Secara visual, tampilan awal dari *dashboard* ini disajikan dalam bentuk Homepage *Dashboard* yang menampilkan branding perusahaan dengan latar aktivitas logistik, serta tiga tombol utama yang dapat diakses pengguna untuk masuk ke masing-masing halaman utama *dashboard*. Tampilan ini dapat dilihat pada Gambar 3.16 Homepage *Dashboard* berikut. Penempatan elemen visual ini juga dirancang agar intuitif dan tetap profesional, sehingga pengguna dapat langsung memahami alur navigasi sejak pertama kali mengakses dashboard.



Gambar 3. 16 Homepage Dashboard

Gambar 3.16 menampilkan *Dashboard* interaktif yang dibangun memiliki tiga halaman utama, yaitu:

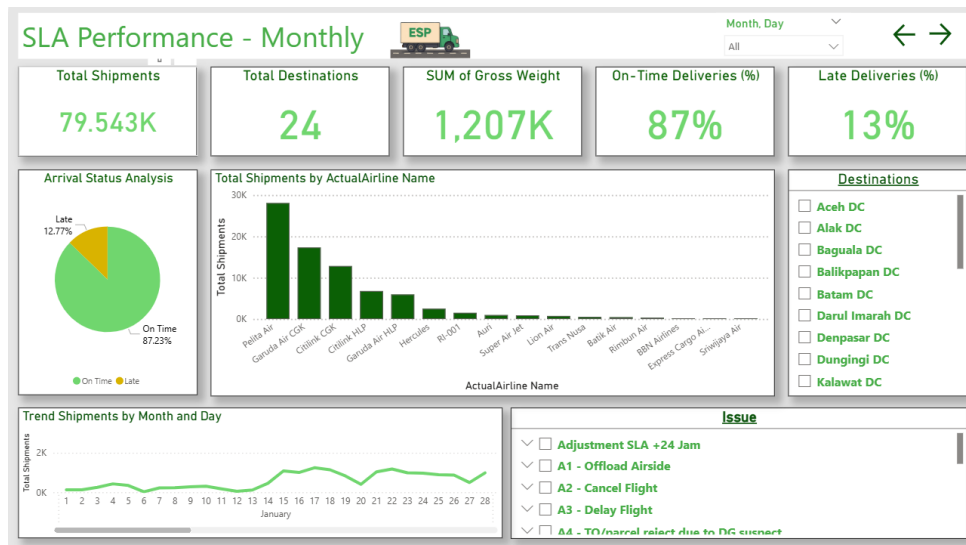
1. ***SLA Performance***, Halaman ini berfokus pada pemantauan kinerja ketepatan waktu pengiriman dengan menampilkan perbandingan antara target *SLA* dan realisasi aktual, serta dilengkapi filter yang memudahkan analisis berdasarkan rute dan periode waktu tertentu.
2. ***Breakdown Delay***, Halaman ini menyajikan rincian penyebab keterlambatan pengiriman, seperti kendala operasional, keterbatasan armada, dan hambatan pada proses distribusi. Visualisasi pada halaman ini membantu dalam mengidentifikasi pola dan faktor dominan penyebab *delay*.
3. ***Financial Overview***, Halaman ini menampilkan ringkasan pendapatan perusahaan yang diperoleh dari aktivitas pengiriman. Informasi yang ditampilkan mencakup total *revenue* dan rata-rata tarif pengiriman per kilogram berdasarkan destinasi, sebagai gambaran umum kinerja finansial.

Pada halaman *overview revenue* disajikan berdasarkan total pengiriman yang dilakukan, dengan mempertimbangkan rate per kilogram

yang berbeda-beda untuk setiap destinasi. Perlu dicatat bahwa halaman ini hanya menyajikan gambaran pendapatan saja, tanpa mengaitkan langsung antara keterlambatan *SLA* dengan dampak finansial, mengingat keterbatasan data keuangan yang tersedia dalam penelitian ini. Ketiga halaman ini saling melengkapi untuk memberikan pandangan menyeluruh terhadap kinerja operasional pengiriman, namun dengan fokus utama tetap diarahkan pada pencapaian *SLA*. Untuk itu, penjelasan selanjutnya akan difokuskan terlebih dahulu pada halaman *SLA Performance* yang menjadi pusat monitoring utama dalam *dashboard* ini.

1. Halaman *SLA Performance*

Halaman pertama dalam *dashboard* interaktif ini adalah *SLA Performance*, yang dirancang sebagai pusat pemantauan utama untuk mengevaluasi sejauh mana pengiriman yang dilakukan perusahaan memenuhi target waktu yang telah disepakati dalam *Service Level Agreement (SLA)*. Pada halaman ini, pengguna dapat secara komprehensif melihat performa pengiriman dalam skala bulanan maupun harian, yang disajikan melalui berbagai indikator kunci dan visualisasi yang saling melengkapi. Tampilan yang dihadirkan sengaja disusun dengan layout yang bersih dan informatif agar dapat diakses dengan mudah oleh berbagai pihak, baik manajemen maupun tim operasional di lapangan. Selain menampilkan metrik *SLA* secara umum, halaman ini juga memungkinkan pengguna untuk melakukan filter berdasarkan destinasi, jenis layanan, serta status keterlambatan, sehingga analisis yang dilakukan dapat lebih tajam dan kontekstual. Grafik tren *SLA* bulanan membantu dalam mengidentifikasi pola performa dalam jangka waktu tertentu, sedangkan *summary card* memberikan gambaran cepat mengenai status pencapaian target. Desain visual yang digunakan mengedepankan warna-warna yang kontras namun profesional untuk mempertegas informasi yang bersifat krusial. Secara visual, halaman *SLA Performance* ini dapat dilihat pada Gambar 3.17 *Halaman SLA Performance* berikut.



Gambar 3. 17 Halaman *SLA Performance*

Gambar 3.17 *Dashboard* ini menyajikan rangkaian metrik performa penting yang langsung terlihat di bagian atas, mulai dari total pengiriman yang telah dilakukan, jumlah destinasi yang dicakup, hingga total berat kiriman yang berhasil dikirimkan selama periode yang dipilih. Selain itu, dua indikator kritikal lainnya, yakni persentase pengiriman tepat waktu dan persentase keterlambatan, juga ditampilkan secara jelas sehingga memudahkan pengguna untuk langsung menangkap kondisi realisasi *SLA* secara keseluruhan. Nilai *on-time deliveries* yang mencapai 87% mencerminkan pencapaian yang masih dalam kategori baik, namun angka *late deliveries* sebesar 13% tetap menjadi perhatian khusus yang memerlukan analisa lebih lanjut. Hal ini juga diperkuat dengan adanya visualisasi *pie chart* di sisi kiri, yang menggambarkan distribusi pengiriman yang berhasil memenuhi *SLA* dan yang mengalami keterlambatan secara proporsional. Lebih dalam lagi, halaman ini menyediakan visualisasi dalam bentuk bar chart yang menampilkan total pengiriman berdasarkan maskapai penerbangan yang digunakan. Dengan adanya grafik ini, pengguna dapat mengetahui maskapai mana yang mendominasi proses pengiriman, serta dapat mengidentifikasi apakah terdapat pola tertentu terkait performa pengiriman antar maskapai. Sebagai contoh, terlihat bahwa maskapai Pelita

Air dan Garuda Air merupakan kontributor utama terhadap total pengiriman, sementara maskapai lainnya seperti Lion Air, Batik Air, dan Rimbun Air memiliki volume pengiriman yang lebih rendah. Informasi ini menjadi penting dalam konteks pengelolaan vendor transportasi dan evaluasi performa penyedia jasa logistik.

Pada bagian bawah halaman, terdapat visualisasi berupa tren pengiriman yang disajikan dalam bentuk *line chart*. Grafik ini memungkinkan pengguna untuk melihat fluktuasi jumlah pengiriman setiap harinya dalam periode satu bulan penuh. Dengan memantau tren ini, pengguna dapat mendeteksi adanya lonjakan pengiriman pada hari-hari tertentu atau justru penurunan yang bisa dikaitkan dengan faktor operasional maupun faktor eksternal seperti cuaca atau regulasi penerbangan. Selain itu, daftar *issue* yang terdapat di bagian bawah *dashboard* memberikan *insight* tambahan mengenai penyebab utama keterlambatan pengiriman. Daftar ini mencakup berbagai kode penyebab seperti pembatalan penerbangan, keterlambatan penerbangan, hingga *proses offload* yang memakan waktu lebih lama di *airside* yang akan disajikan secara lengkap pada Tabel 3.4 berikut.

Tabel 3. 5 Code & Issue

Code	Issue
A1	<i>Offload Airside</i>
A2	<i>Cancel Flight</i>
A3	<i>Delay Flight</i>
A4	<i>TO/parcel reject due to DG suspect</i>
A5	<i>Partial Flight</i>
A6	<i>Flight Transit (destination forwarded area)</i>
A7	<i>No Space for Booking (Normal Days)</i>
A8	<i>No Space for Booking (Peak Season)</i>
A9	<i>Limitation of Airport operations schedule</i>
A13	<i>Delay Ready to Dispatch</i>
D12	<i>Slot Doorring</i>

Code	Issue
D14	<i>Late Depart (Lini 2) due to Vendor Resources</i>
D15	<i>Late Doorring (By Request Destination)</i>
D16	<i>Unloading Issue (Slot Doorring and Multidrop destination hub)</i>
D17	<i>Road Restrictions</i>
D18	<i>Overload Facility SPX</i>
D19	<i>Late Doorring By Vendor (Vehicle Issue / Accident)</i>
O10	<i>Sorting and Booking Process (Delay)</i>
O11	<i>Late Arrival at Crossdock Warehouse (Origin)</i>

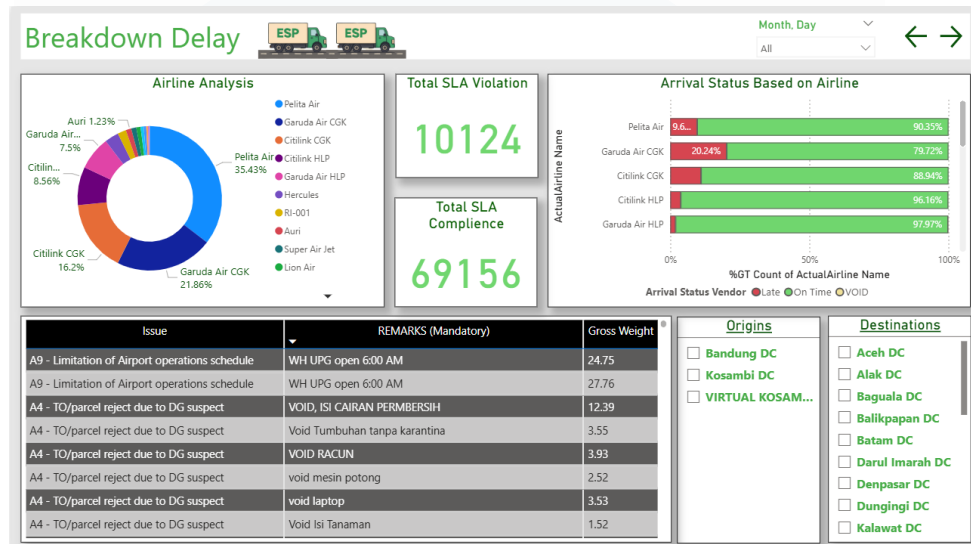
Table 3.4 yang mencakup berbagai penyebab keterlambatan yang menjadi salah satu fokus penting di dalam alur pengiriman barang selain itu, fitur interaktif lainnya yang disematkan dalam halaman ini adalah filter yang memungkinkan pengguna untuk melakukan analisa lebih mendalam. Pengguna dapat memilih periode waktu tertentu, baik bulanan maupun harian, serta menyaring data berdasarkan destinasi tujuan pengiriman. Dengan memilih destinasi tertentu, misalnya Aceh DC atau Batam DC, maka seluruh visualisasi yang ada akan otomatis menyesuaikan sehingga hanya menampilkan data yang relevan dengan filter yang dipilih. Keunggulan ini memberikan fleksibilitas tinggi dalam analisa, terutama bagi tim operasional yang ingin melakukan evaluasi performa pengiriman di wilayah tertentu. Secara keseluruhan, halaman *SLA Performance* ini berperan penting sebagai alat kontrol dan monitoring yang membantu pihak manajemen untuk memastikan bahwa proses pengiriman berjalan sesuai standar layanan yang telah ditetapkan. Dengan tampilan yang menyajikan data secara berkala dan interaktif, pengguna dapat dengan cepat mengidentifikasi area yang membutuhkan perbaikan, mengambil keputusan berbasis data, serta mengoptimalkan proses pengiriman untuk meningkatkan kepuasan pelanggan dan efisiensi operasional perusahaan.

Setelah memahami performa *SLA* secara menyeluruh di halaman ini, pembahasan selanjutnya akan diarahkan ke halaman kedua dalam *dashboard*, yaitu halaman *Breakdown Delay*, yang secara khusus menyajikan analisa lebih rinci mengenai penyebab keterlambatan pengiriman serta distribusinya ke setiap destinasi yang terlibat.

2. Halaman *Breakdown Delay*

Halaman kedua dari *dashboard* ini, yaitu *Breakdown Delay*, dirancang khusus untuk memberikan analisa yang lebih dalam dan komprehensif terkait dengan pelanggaran *SLA* yang terjadi dalam proses pengiriman. Jika pada halaman sebelumnya fokus utamanya adalah memantau performa secara agregat dan umum, maka halaman ini berfungsi sebagai alat diagnostik yang lebih rinci untuk mengurai dan mengidentifikasi lebih lanjut berbagai sumber keterlambatan yang secara langsung memengaruhi kinerja pengiriman. Keberadaan halaman ini menjadi sangat krusial karena membantu manajemen dan tim operasional tidak hanya dalam mengetahui seberapa besar pelanggaran *SLA* yang terjadi, tetapi juga dalam menelusuri maskapai yang terlibat, rute tujuan yang terdampak, serta penyebab spesifik yang berkontribusi terhadap keterlambatan tersebut, seperti kendala teknis, operasional, atau bahkan faktor eksternal. Dengan demikian, halaman *Breakdown Delay* menjadi pondasi utama yang kuat dalam upaya *continuous improvement* terhadap sistem logistik perusahaan secara keseluruhan. Visualisasi dan fitur analitik yang ditampilkan dalam halaman ini memungkinkan pengambilan keputusan yang lebih tepat dan berbasis data. Selain itu, halaman ini dilengkapi dengan fitur interaktif yang memungkinkan pengguna memfilter data berdasarkan parameter seperti tanggal, origin, atau maskapai, sehingga analisis dapat lebih terarah. Fitur ini membantu mempercepat identifikasi akar masalah dan mendukung pengambilan keputusan yang lebih responsif terhadap kondisi operasional. Secara visual, tampilan dari halaman

Breakdown Delay ini dapat dilihat pada Gambar 3.18 Halaman *Breakdown Delay* berikut.



Gambar 3. 18 Halaman Breakdown Delay

Gambar 3.18 *Dashboard* ini menampilkan beberapa visual utama yang saling melengkapi dalam proses analisa. Pada bagian kiri atas, terdapat *donut chart* dengan judul *Airline Analysis* yang menggambarkan distribusi pelanggaran *SLA* berdasarkan maskapai penerbangan yang digunakan. Dari visual ini, terlihat bahwa maskapai Pelita Air memegang porsi terbesar, yaitu 35,43% dari total pelanggaran, disusul oleh Garuda Air CGK (21,86%) dan Citilink CGK (16,2%). Informasi ini memungkinkan perusahaan untuk mengevaluasi performa vendor transportasi secara lebih objektif, serta menjadi dasar dalam menyusun strategi negosiasi atau penyesuaian alokasi volume pengiriman di masa mendatang. Tepat di sebelahnya, dua indikator kunci menampilkan angka *Total SLA Violation* sebesar 10.124 dan *Total SLA Compliance* sebesar 69.156, yang memberikan gambaran umum mengenai rasio pelanggaran terhadap kepatuhan *SLA* secara keseluruhan. Di sisi kanan atas, terdapat grafik horizontal berjudul *Arrival Status Based on Airline* yang memvisualisasikan persentase pengiriman yang tiba tepat waktu (*on time*), terlambat (*late*), atau *void* untuk masing-masing maskapai. Data ini menunjukkan performa yang

cukup beragam, di mana maskapai seperti Garuda Air HLP mencatat tingkat ketepatan tertinggi dengan 97,97% pengiriman *on time*, sementara Garuda Air CGK justru memiliki catatan ketepatan paling rendah, yaitu hanya 79,72%. Visualisasi ini mempermudah tim operasional dalam melakukan *benchmarking* antar maskapai, serta memetakan risiko keterlambatan yang mungkin terjadi berdasarkan pilihan vendor pengangkutan.

Pada bagian bawah *dashboard*, terdapat tabel yang menampilkan daftar *issues* secara rinci, lengkap dengan keterangan *remarks* yang wajib diisi dan data *gross weight* pengiriman yang terdampak. Tabel ini memberikan lapisan informasi tambahan yang sangat berguna untuk keperluan *root cause analysis*. Misalnya, terdapat isu "A9 - *Limitation of Airport operations schedule*" yang beberapa kali muncul sebagai penyebab keterlambatan dengan catatan bahwa *warehouse* (WH) di UPG baru dibuka pukul 6 pagi. Selain itu, terdapat pula data *void* yang menunjukkan pengiriman yang ditolak karena isi paket mencurigakan, seperti cairan pembersih, tanaman tanpa karantina, atau *laptop* yang dikategorikan berisiko. Dengan adanya tabel ini, tim dapat lebih mudah melacak penyebab utama keterlambatan dan melakukan tindakan korektif yang lebih tepat sasaran.

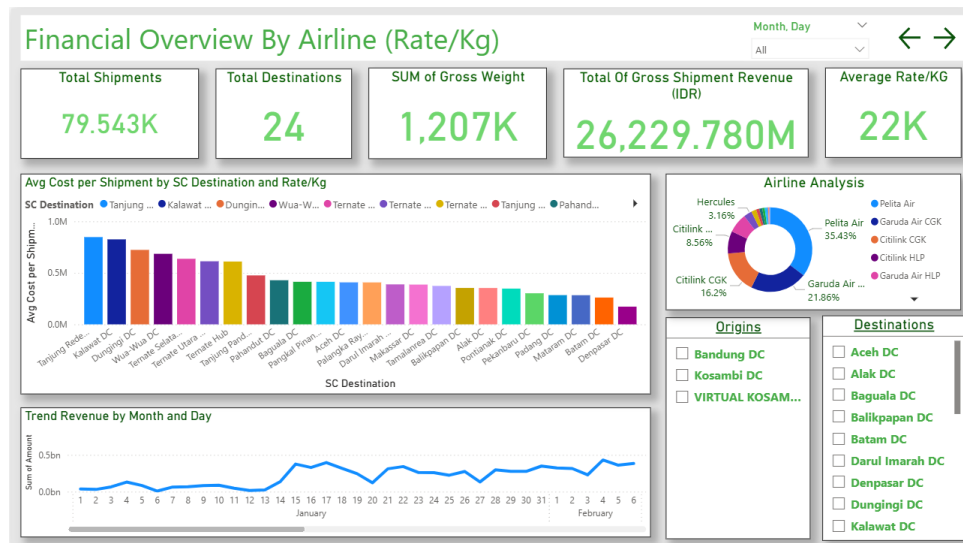
Fitur interaktif yang tersedia dalam halaman ini juga tak kalah pentingnya. Pengguna dapat menggunakan filter untuk memilih data berdasarkan bulan, hari, asal pengiriman (*origins*), maupun tujuan pengiriman (*destinations*). Filter ini memungkinkan *dashboard* berfungsi secara dinamis, di mana seluruh visual akan otomatis menyesuaikan tampilan datanya sesuai filter yang dipilih. Dengan demikian, pengguna dapat melakukan analisa mendalam untuk memantau performa *SLA* secara spesifik di wilayah tertentu, misalnya memeriksa pengiriman dari Bandung DC menuju Batam DC, atau memfilter hanya pengiriman yang terjadi pada periode *peak season*. Fleksibilitas ini memberikan manfaat besar bagi tim

dalam merancang strategi perbaikan *SLA* secara lebih terarah dan kontekstual.

Secara keseluruhan, halaman *Breakdown Delay* ini berperan sebagai *diagnostic tool* yang membantu perusahaan dalam menggali akar masalah pelanggaran *SLA*, serta sebagai acuan dalam perumusan langkah-langkah peningkatan performa logistik di masa yang akan datang. Dengan visualisasi yang informatif dan fitur yang interaktif, *dashboard* ini menjadi alat bantu yang esensial dalam menjaga standar layanan tetap tinggi dan operasional distribusi tetap optimal. Setelah melakukan analisa mendalam terkait performa pengiriman dan faktor penyebab keterlambatan, *dashboard* kemudian mengarahkan pengguna menuju halaman berikutnya, yaitu *Financial Overview*, yang akan menyajikan dimensi keuangan dari seluruh proses distribusi ini, termasuk tren revenue dan biaya operasional yang berkaitan langsung dengan performa *SLA* dan efisiensi logistik.

3. Halaman *Financial Overview*

Halaman terakhir yaitu *Financial Overview* berfungsi sebagai pusat informasi yang menampilkan performa keuangan secara komprehensif dari seluruh aktivitas pengiriman yang dilakukan oleh perusahaan. Jika halaman sebelumnya berfokus pada aspek operasional dan kendala pengiriman, maka halaman ini lebih menyoroti bagaimana aktivitas tersebut berdampak langsung terhadap kinerja finansial. Tampilan ini menyajikan metrik-metrik penting seperti total jumlah pengiriman, total destinasi yang dilayani, total berat kotor kiriman (*gross weight*), hingga total pendapatan kotor dari pengiriman yang diukur dalam satuan Rupiah. Selain itu, terdapat indikator yang menampilkan rata-rata tarif pengiriman barang per kilogram (*Average Rate/Kg*), yang nilainya mencapai 22 ribu rupiah. Angka ini menjadi parameter penting dalam mengevaluasi struktur biaya yang dikenakan kepada pelanggan dan membantu perusahaan dalam memantau sejauh mana harga jual per kilo mampu mengimbangi biaya logistik yang dikeluarkan yang rinciannya akan ditampilkan pada Gambar 3.19.



Gambar 3. 19 Halaman Financial Overview

Gambar 3.19 pada tampilan halaman ini, data disajikan secara lebih mendetail melalui grafik *Avg Cost per Shipment* yang mengelompokkan biaya rata-rata per pengiriman berdasarkan tujuan distribusi (*SC Destination*). Grafik ini memudahkan identifikasi destinasi yang memiliki rata-rata biaya tertinggi, seperti Tanjung Redeb, Kalawat DC, dan Dungingi DC, yang masing-masing menunjukkan biaya mendekati angka 1 juta rupiah per pengiriman. Data semacam ini sangat penting untuk pengambilan keputusan dalam hal penyesuaian tarif pengiriman maupun efisiensi jalur distribusi. Di bagian bawah, terdapat grafik tren pendapatan (*Trend Revenue*) yang menampilkan fluktuasi pemasukan perusahaan secara harian, memberikan pandangan jangka waktu tertentu untuk melihat pola pendapatan yang mungkin berkaitan dengan musim puncak maupun kendala operasional yang terjadi sebelumnya. Tidak hanya itu, *pie chart* pada sisi kanan halaman kembali menampilkan distribusi penggunaan maskapai, yang memberi konteks tambahan terkait hubungan antara performa maskapai dengan tingkat biaya dan pendapatan yang dihasilkan. Fitur interaktif seperti filter asal pengiriman (*Origins*) dan destinasi pengiriman (*Destinations*) tetap dipertahankan, sehingga memungkinkan pengguna untuk mengkustomisasi tampilan data berdasarkan lokasi

tertentu. Dengan desain yang informatif dan interaktif ini, halaman *Financial Overview* tidak hanya mendukung fungsi pelaporan keuangan, tetapi juga menjadi alat bantu dalam penyusunan strategi bisnis yang berbasis data.

3.3 Kendala yang Ditemukan

Sepanjang periode magang di PT Eka Satya Puspita, aktivitas secara umum berlangsung dengan baik tanpa hambatan besar. Namun, terdapat sejumlah kendala teknis dan operasional yang cukup menantang selama pelaksanaan tugas, khususnya dalam pengelolaan dan analisis data. Kendala-kendala ini dapat diuraikan sebagai berikut:

1. **Volume data yang sangat besar dan dinamis.**

Setiap harinya, perusahaan menangani ratusan hingga ribuan pengiriman ke berbagai daerah di Indonesia. Tingginya intensitas pengiriman ini menghasilkan data dalam jumlah besar yang sangat dinamis dan terus berubah. Kondisi ini membuat proses pelacakan performa pengiriman, seperti *SLA (Service Level Agreement)*, menjadi kompleks. Ketelitian sangat dibutuhkan dalam menyaring data yang relevan, serta menghindari kesalahan interpretasi akibat data yang belum diperbarui secara real-time. Selain itu, perubahan data harian ini turut memengaruhi kesesuaian antar laporan dan mengharuskan peninjauan berkala.

2. **Keterbatasan dalam pemanfaatan alat analisis yang tersedia.**

Dalam operasional sehari-hari, perusahaan menggunakan Microsoft Excel sebagai alat utama untuk pelaporan dan evaluasi data logistik, termasuk melalui fitur *pivot table* dan grafik sederhana. Meskipun Excel memiliki kemampuan analisis data yang cukup kuat, pemanfaatannya di perusahaan belum maksimal. Hal ini disebabkan oleh keterbatasan pemahaman terhadap fitur-fitur lanjutan serta kompleksitas penggunaan rumus-rumus tertentu, terutama saat menangani data dalam volume besar. Akibatnya, proses penyusunan laporan menjadi kurang efisien dan kurang mendalam, khususnya ketika dibutuhkan dalam waktu singkat untuk *daily meeting* atau

presentasi kepada manajemen. Dampaknya, informasi yang disajikan belum mampu menggambarkan kondisi logistik secara komprehensif, sehingga menyulitkan proses evaluasi dan pengambilan keputusan secara menyeluruh.

3. **Kompleksitas dalam penyajian dan interpretasi data.**

Proses penyusunan laporan dari data mentah sering kali membutuhkan waktu lebih karena data tidak langsung dapat digunakan. Data yang disusun secara manual perlu disesuaikan ulang agar dapat dipahami oleh seluruh pemangku kepentingan lintas divisi. Akibatnya, proses evaluasi kinerja menjadi terhambat, dan pengambilan keputusan strategis pun mengalami keterlambatan karena harus menunggu data dirapikan terlebih dahulu.

4. **Kendala teknis dalam pembuatan model clustering.**

Penerapan algoritma *K-Means Clustering* untuk analisis pengelompokan data menghadapi hambatan teknis, terutama terkait keterbatasan spesifikasi perangkat yang digunakan. Komputasi dengan dataset besar menyebabkan proses berjalan lambat bahkan gagal dieksekusi secara penuh. Untuk mengatasi hal ini, dilakukan pendekatan *sampling*, namun hal ini turut mengurangi presisi analisis karena tidak semua variasi data terwakili. Kendala ini juga menimbulkan kebutuhan untuk menyesuaikan skala proyek dengan kemampuan komputasi yang tersedia.

5. **Struktur data yang kompleks dan tidak seragam.**

Dataset pengiriman yang digunakan mengandung banyak elemen seperti kode status pengiriman, jenis keterlambatan, serta variasi rute berdasarkan asal dan tujuan. Setiap elemen ini memiliki nilai dan format yang berbeda, sehingga proses *data cleaning* menjadi tahap yang memakan waktu dan memerlukan ketelitian tinggi. Kesalahan dalam interpretasi satu kode atau status saja dapat menyebabkan ketidaksesuaian analisis pada keseluruhan dataset. Proses ini menjadi semakin menantang ketika tidak tersedia dokumentasi data yang lengkap dari pihak operasional.

6. **Kendala dalam integrasi antar divisi.**

Setiap divisi seperti operasional, vendor, dan administrasi terkadang memiliki format pencatatan yang berbeda, baik dari segi struktur data, istilah, hingga frekuensi pembaruan. Ketika data dari beberapa divisi harus dikonsolidasikan untuk keperluan analisis, proses penyamaan format dan pemahaman menjadi tantangan tersendiri. Hal ini diperparah oleh belum adanya sistem integrasi terpadu antar divisi, sehingga koordinasi data masih dilakukan secara manual atau semi-otomatis, dengan potensi miskomunikasi cukup tinggi. Salah satu contohnya adalah perbedaan definisi keterlambatan atau status *SLA* antar tim.

7. **Keterbatasan dalam kapasitas analitik sumber daya manusia.**

Tim internal sebagian besar masih terbiasa menggunakan alat-alat konvensional dan belum terbiasa dengan penggunaan teknologi analitik yang lebih modern, seperti Power BI, Google Looker Studio, atau pemodelan dengan Python. Hal ini menyebabkan adanya hambatan dalam proses adaptasi terhadap alat dan metode baru yang digunakan dalam proyek magang. Akibatnya, meskipun solusi sudah tersedia secara teknis, implementasinya memerlukan waktu lebih karena harus melalui proses pelatihan atau pendampingan terlebih dahulu.

Secara keseluruhan, kendala-kendala di atas menunjukkan bahwa kompleksitas operasional di industri logistik seperti PT Eka Satya Puspita membutuhkan pendekatan analisis data yang tidak hanya akurat, tetapi juga fleksibel dan mampu menjembatani kesenjangan antar tim serta sistem. Volume data yang besar, keterbatasan alat, serta tantangan integrasi menjadi faktor dominan yang memengaruhi efektivitas pelaporan dan pengambilan keputusan selama masa magang berlangsung.

3.4 Solusi atas Kendala yang Ditemukan

Dalam menjalani masa magang di PT Eka Satya Puspita, terdapat sejumlah tantangan yang harus dihadapi sepanjang prosesnya, sejumlah solusi telah dirancang dan diterapkan guna mengatasi tantangan baik dari sisi teknis, proses

bisnis, maupun penyajian data. Solusi-solusi tersebut diimplementasikan secara bertahap dan berfokus pada peningkatan efisiensi, akurasi, serta kolaborasi lintas tim. Beberapa solusi utama yang dijalankan meliputi:

1. **Pengembangan dashboard interaktif sebagai pengganti analisis manual.**

Sebagai upaya meningkatkan efisiensi analisis data, dikembangkan sebuah dashboard interaktif yang mampu menangani volume data pengiriman yang besar dan dinamis. Dashboard ini menyajikan metrik penting seperti performa *SLA* dan tingkat keterlambatan dalam bentuk visualisasi yang mudah dipahami. Hasilnya, proses *daily meeting* yang sebelumnya lambat dan bersifat statis kini dapat dilakukan secara lebih cepat, interaktif, dan berbasis data real-time.

2. **Penerapan teknik sampling untuk mengatasi keterbatasan memori.**

Dalam proses pemodelan K-Means yang memerlukan komputasi tinggi, diterapkan strategi pengambilan *sample* data yang representatif terhadap keseluruhan populasi pengiriman. Meskipun terdapat sedikit penurunan akurasi, pendekatan ini tetap memungkinkan analisis risiko logistik dilakukan dengan hasil yang cukup valid tanpa membebani kapasitas perangkat.

3. **Penyusunan ulang pipeline *data cleaning* dan transformasi.**

Proses pembersihan dan pemetaan data yang semula memakan waktu diperbaiki dengan membangun pipeline pengolahan data yang lebih terstruktur dan sistematis. Hal ini mempercepat proses pada tahap analisis lanjutan, sekaligus mengurangi potensi kesalahan saat dilakukan visualisasi atau pemodelan.

4. **Peningkatan integrasi data antar divisi.**

Dilakukan penyelarasan format dan definisi antara tim operasional, vendor, dan administrasi terkait pencatatan *SLA* dan penyebab keterlambatan. Komunikasi lintas tim juga ditingkatkan melalui diskusi terbuka untuk

memperjelas kendala operasional di lapangan, seperti keterlambatan di jalur udara dan darat. Hasilnya, proses konsolidasi data menjadi lebih efisien dan miskomunikasi dapat diminimalisir.

5. **Peningkatan kapasitas analitik sumber daya manusia.**

Diberikan sesi *knowledge sharing* terkait penggunaan dashboard interaktif agar tim internal dapat mulai beradaptasi dengan teknologi analitik modern. Fitur-fitur seperti filter data, segmentasi destinasi, dan analisis tren waktu dimanfaatkan untuk mendorong kemandirian tim dalam menyusun laporan maupun menganalisis performa pengiriman.

Secara keseluruhan, solusi-solusi yang diterapkan selama masa magang ini difokuskan pada dua pendekatan utama: optimalisasi teknologi dan penyelarasan proses lintas divisi. Sinergi dari keduanya berkontribusi langsung terhadap peningkatan efisiensi operasional serta ketepatan pengambilan keputusan berbasis data di lingkungan PT Eka Satya Puspita.