

BAB 1

PENDAHULUAN

Perkembangan teknologi informasi yang semakin pesat mendorong berbagai organisasi untuk mengoptimalkan pengelolaan data sebagai sumber daya strategis [1]. Pemanfaatan data tidak hanya terbatas pada penyimpanan, tetapi juga mencakup proses pengolahan, transformasi, hingga penyajian data secara otomatis dengan visualisasi data yang ringkas untuk memantau *key performance index* (KPI), *trend* penjualan dan menentukan keputusan bisnis [1]. Hal tersebut menuntut adanya sistem *pipeline data* yang terstruktur, andal, serta mampu berjalan secara otomatis dan konsisten setiap periode waktu tertentu. Pada konteks tersebut, dibutuhkan integrasi berbagai layanan berbasis *cloud* yang mampu mendukung proses *extract*, transformasi, *load*, orkestrasi, dan visualisasi data secara efisien.

1.1 Latar Belakang Masalah

Banyak perusahaan menghadapi tantangan dalam mengelola aliran data dari berbagai sumber secara terpusat dan terotomatisasi [2]. Proses *ingestion data* yang dilakukan secara manual berpotensi menimbulkan inkonsistensi, keterlambatan pemrosesan, serta risiko kesalahan input. Selain itu, data mentah yang tersimpan tanpa proses transformasi tidak dapat langsung dimanfaatkan untuk kebutuhan analisis [3], sehingga diperlukan mekanisme pengolahan data yang sistematis agar data memiliki struktur yang sesuai dengan kebutuhan visualisasi dan analitik.

Ketika volume data semakin meningkat dan kebutuhan pelaporan bersifat harian atau periodik, sistem manual menjadi tidak efisien untuk diterapkan dan tidak memiliki makna jika tidak diolah [3]. Diperlukan sistem *pipeline data* yang mampu berjalan secara otomatis mulai dari proses pengambilan data, penyimpanan ke dalam gudang data (*data warehouse*), transformasi menggunakan alat yang konsisten dan terstandarisasi, hingga penyajian data untuk keperluan visualisasi melalui platform analitik seperti Looker Studio ataupun Google Spreadsheets.

Google Cloud Platform menyediakan serangkaian layanan yang dapat diintegrasikan untuk membangun *pipeline* data yang otomatis dan *scalable* tanpa memerlukan pengelolaan infrastruktur yang besar [4]. Google Workflows berfungsi sebagai mesin orkestrasi *serverless* yang dapat memanggil layanan lain seperti Cloud Run Jobs atau BigQuery, dengan model biaya *pay-per-use* yang hanya

menagih ketika workflow dieksekusi, membuatnya efisien untuk alur data ringan sampai menengah. Cloud Scheduler memungkinkan penjadwalan tugas secara periodik dengan biaya minimal per *job* yang didefinisikan. Pendekatan ini dapat mengeliminasi kebutuhan akan platform orkestrasi dengan biaya tinggi seperti Airflow dikarenakan memerlukan server, terutama untuk *pipeline* yang tidak memerlukan logika orkestrasi tingkat lanjut atau dependensi kompleks antara tugas.

Berdasarkan kondisi tersebut, dirancang dan diimplementasikan sebuah *pipeline data* otomatis berbasis GCP di PT Entrefine Data Indonesia untuk mengatasi tantangan proses *ingestion*, transformasi, orkestrasi, dan penjadwalan data secara berkelanjutan. Sistem ini diharapkan mampu meningkatkan efisiensi operasional, menjaga konsistensi data, serta mendukung proses visualisasi data melalui platform analitik seperti Looker Studio dan Google Spreadsheets.

1.2 Maksud dan Tujuan Kerja Magang

Maksud pelaksanaan kerja magang di PT Entrefine Data Indonesia adalah sebagai berikut:

1. Memberikan pemahaman langsung mengenai praktik profesional dalam bidang pengelolaan data di lingkungan industri
2. Memperluas wawasan teknis serta meningkatkan kemampuan adaptasi terhadap sistem kerja berbasis proyek

Tujuan dari pelaksanaan kerja magang ini adalah merancang dan mengimplementasikan *pipeline data* otomatis berbasis Google Cloud Platform (GCP) untuk proses pengolahan data di PT Entrefine Data Indonesia.

1.3 Waktu dan Prosedur Pelaksanaan Kerja Magang

Kegiatan kerja magang dilaksanakan mulai 1 Agustus 2025 hingga 31 Januari 2026 dengan total durasi selama enam bulan. Aktivitas kerja berlangsung setiap Senin hingga Jumat pada rentang waktu 09.30 hingga 17.00 WIB. Seluruh kegiatan dilakukan secara *onsite* di kantor PT Entrefine Data Indonesia yang beralamat di Jl. Edutown No.A5, Pagedangan, Kabupaten Tangerang, Banten 15339(Piazza Mozia BSD).

Prosedur pelaksanaan kerja mengikuti pola operasional harian yang telah ditetapkan perusahaan. Setiap hari dimulai dengan sesi *briefing* untuk menyusun

prioritas pekerjaan dan memastikan keselarasan proses kerja dengan kebutuhan proyek. Setelah *briefing*, tugas diberikan oleh supervisor untuk kemudian dieksekusi sesuai alur teknis yang telah dijelaskan. Hasil pekerjaan akan melalui proses verifikasi untuk memastikan kesesuaian dengan standar implementasi yang berlaku di lingkungan perusahaan. Jika ditemukan ketidaksesuaian, dilakukan penyempurnaan hingga memenuhi kriteria yang ditetapkan.

Pengelolaan alur tugas dan dokumentasi progres kerja dilakukan melalui Trello sebagai media koordinasi internal. Selain *briefing* pagi, sesi evaluasi singkat juga dilakukan pada akhir hari kerja untuk meninjau hasil pelaksanaan tugas dan menentukan fokus pekerjaan berikutnya. Seluruh aktivitas yang dijalankan terintegrasi langsung dengan proyek operasional perusahaan, sehingga setiap proses memiliki dampak implementatif terhadap sistem yang sedang dikembangkan.

Supervisor perusahaan berperan sebagai penanggung jawab teknis yang memfasilitasi pengawasan alur kerja. Mekanisme pembimbingan diterapkan secara langsung melalui pemberian tugas berbasis kebutuhan proyek, sedangkan bantuan teknis diberikan apabila terdapat hambatan dalam proses penyelesaian pekerjaan. Dengan prosedur tersebut, kegiatan kerja magang berjalan secara sistematis, terarah, dan sesuai dengan standar kerja profesional di lingkungan industri.

