

PRAKTIKUM REKAYASA PERANGKAT LUNAK 2



PROPOSAL PENAWARAN PENGEMBANGAN PERANGKAT LUNAK

“Pengembangan Automated Data Pipeline Data Cuaca Menggunakan Weatherstack API, Apache Airflow, PostgreSQL, dbt, Apache Superset, dan Docker”

Ahmad Maulana

50422124

Alya Gustasya

50422186

Kelas : 4IA16

Jurusan : Informatika

UNIVERSITAS GUNADARMA

FAKULTAS TEKNOLOGI INDUSTRI

2025

1. Pendahuluan

1.1 Latar Belakang

Perkembangan teknologi informasi mendorong organisasi untuk mengolah data secara cepat, akurat, dan berkelanjutan. Salah satu jenis data yang banyak dimanfaatkan dalam berbagai sektor seperti transportasi, pertanian, logistik, dan perencanaan wilayah adalah data cuaca. Data cuaca umumnya disediakan oleh penyedia layanan dalam bentuk Application Programming Interface (API), sehingga diperlukan sistem yang mampu mengambil, menyimpan, mengolah, dan menyajikan data tersebut secara otomatis.

Namun, pengambilan data secara manual atau tidak terjadwal berpotensi menimbulkan ketidakkonsistenan data dan menyulitkan proses analisis. Oleh karena itu, dibutuhkan sebuah sistem **Automated Data Pipeline** yang mampu melakukan proses *Extract, Load, dan Transform (ELT)* secara terjadwal dan terintegrasi.

Proyek ini mengusulkan pengembangan Automated Data Pipeline berbasis Weatherstack API dengan memanfaatkan Apache Airflow sebagai orchestrator, PostgreSQL sebagai database, dbt untuk transformasi data, Apache Superset untuk visualisasi data, serta Docker untuk containerization sistem.

2. Gambaran Umum Sistem

Sistem yang dikembangkan merupakan pipeline data otomatis yang bekerja secara end-to-end. Sistem ini akan mengambil data cuaca dari Weatherstack API secara berkala, menyimpannya ke dalam database PostgreSQL, melakukan transformasi data analitik, dan menampilkan hasilnya dalam bentuk dashboard interaktif.

2.1 Deskripsi Sistem

Alur kerja sistem secara umum adalah sebagai berikut:

1. Sistem mengambil data cuaca dari Weatherstack API menggunakan script Python.
2. Data mentah (raw data) disimpan ke dalam database PostgreSQL.
3. Apache Airflow mengatur penjadwalan dan orkestrasi seluruh proses pipeline.

4. dbt digunakan untuk melakukan transformasi data menjadi data analitik.
 5. Apache Superset menampilkan data dalam bentuk dashboard visual.
 6. Seluruh layanan dijalankan dalam container Docker.
-

3. Ruang Lingkup Pengembangan

Ruang lingkup pengembangan perangkat lunak pada proyek ini meliputi:

- Integrasi Weatherstack API sebagai sumber data cuaca
- Pengambilan data cuaca secara otomatis dan terjadwal
- Penyimpanan data ke database PostgreSQL
- Transformasi data menggunakan dbt
- Orkestrasi workflow menggunakan Apache Airflow
- Visualisasi data menggunakan Apache Superset
- Containerization sistem menggunakan Docker dan Docker Compose

Adapun batasan sistem dalam proyek ini adalah:

- Data yang diambil terbatas pada data cuaca dari Weatherstack API
 - Sistem berjalan pada lingkungan lokal berbasis Docker
 - Fokus sistem pada proses data engineering, bukan prediksi cuaca
-

4. Metode Pengembangan Perangkat Lunak

Metode pengembangan perangkat lunak yang digunakan adalah **Metode Waterfall**.

Metode ini dipilih karena kebutuhan sistem telah didefinisikan dengan jelas sejak awal dan tahapan pengembangan dilakukan secara berurutan.

4.1 Tahapan Metode Waterfall

1. Analisis	Kebutuhan
Mengidentifikasi kebutuhan fungsional dan non-fungsional sistem, termasuk kebutuhan data, tools, dan output sistem.	
2. Perancangan	Sistem
Merancang arsitektur data pipeline, alur ELT, serta hubungan antar komponen sistem.	
3. Implementasi	
Mengimplementasikan script Python untuk ingestion data, konfigurasi Airflow DAG, model dbt, dan Docker Compose.	
4. Pengujian	
Melakukan pengujian terhadap setiap komponen sistem dan pengujian end-to-end pipeline.	
5. Deployment	dan
Dokumentasi	
Menjalankan sistem secara penuh dan menyusun dokumentasi penggunaan sistem.	

5. Timeline Pengerjaan

No	Waktu	Kegiatan
1	Minggu 1	Analisis kebutuhan dan setup environment (WSL, Docker, VS Code)
2	Minggu 2	Integrasi Weatherstack API dan ingestion data ke PostgreSQL
3	Minggu 3	Setup Apache Airflow dan pembuatan DAG ingestion
4	Minggu 4	Implementasi dbt dan transformasi data
5	Minggu 5	Integrasi Airflow dengan dbt dan dependency DAG
6	Minggu 6	Setup Apache Superset dan pembuatan dashboard
7	Minggu 7	Pengujian sistem end-to-end
8	Minggu 8	Penyusunan dokumentasi dan laporan akhir

6. Penutup

Dengan adanya proposal ini, diharapkan pengembangan Automated Data Pipeline dapat berjalan secara terstruktur dan sesuai dengan kebutuhan. Sistem yang dikembangkan diharapkan mampu menjadi contoh implementasi data engineering end-to-end serta memberikan manfaat sebagai media pembelajaran dan portfolio proyek.