

PRAKTIKUM REKAYASA PERANGKAT LUNAK 2



LAPORAN PENDAHULUAN

**“Pengembangan Automated Data Pipeline Data Cuaca Menggunakan Weatherstack
API Berbasis Apache Airflow dan Docker”**

Ahmad Maulana

50422124

Alya Gustasya

50422186

Kelas : 4IA16

Jurusan : Informatika

UNIVERSITAS GUNADARMA

FAKULTAS TEKNOLOGI INDUSTRI

2025

1. Latar Belakang

Perkembangan teknologi informasi dan meningkatnya kebutuhan akan data yang akurat dan real-time mendorong organisasi untuk membangun sistem pengolahan data yang terotomatisasi. Data tidak lagi diolah secara manual, melainkan melalui pipeline yang terstruktur agar dapat digunakan secara efektif untuk analisis dan pengambilan keputusan.

Salah satu jenis data yang memiliki peran penting dalam berbagai sektor adalah data cuaca. Data cuaca digunakan dalam bidang transportasi, pertanian, logistik, mitigasi bencana, serta perencanaan aktivitas harian. Umumnya, data cuaca disediakan oleh pihak ketiga melalui layanan Application Programming Interface (API), seperti Weatherstack.

Permasalahan yang sering muncul adalah data yang diambil secara manual atau tidak terjadwal berpotensi tidak konsisten, sulit ditelusuri, dan tidak siap untuk analisis lanjutan. Selain itu, tanpa adanya orkestrasi yang baik, proses pengambilan, penyimpanan, dan transformasi data menjadi tidak efisien.

Oleh karena itu, diperlukan sebuah sistem **Automated Data Pipeline** yang mampu mengelola data cuaca secara otomatis mulai dari proses pengambilan data, penyimpanan, transformasi, hingga visualisasi. Dengan memanfaatkan Apache Airflow sebagai orchestrator, PostgreSQL sebagai basis data, dbt sebagai alat transformasi data, Apache Superset sebagai media visualisasi, serta Docker sebagai teknologi containerization, sistem ini diharapkan mampu memberikan solusi pengolahan data yang terstruktur dan scalable.

2. Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dapat diidentifikasi beberapa permasalahan sebagai berikut:

1. Data cuaca dari API belum dikelola secara otomatis dan terjadwal.
2. Proses pengolahan data cuaca masih belum terintegrasi dalam satu sistem.
3. Data mentah belum siap digunakan untuk analisis tanpa proses transformasi.
4. Belum tersedia dashboard visual untuk memantau data cuaca secara interaktif.

3. Rumusan Masalah

Rumusan masalah dalam proyek ini adalah sebagai berikut:

1. Bagaimana membangun automated data pipeline untuk pengambilan data cuaca dari Weatherstack API?

2. Bagaimana mengorkestrasi proses pengambilan dan pengolahan data menggunakan Apache Airflow?
3. Bagaimana melakukan transformasi data agar siap untuk analisis menggunakan dbt?
4. Bagaimana menyajikan data cuaca dalam bentuk dashboard menggunakan Apache Superset?

4. Tujuan Penelitian

Tujuan dari pengembangan sistem ini adalah:

1. Membangun sistem automated data pipeline berbasis API.
2. Mengimplementasikan Apache Airflow sebagai sistem orkestrasi dan penjadwalan data.
3. Mengolah dan mentransformasikan data cuaca menggunakan dbt.
4. Menyajikan data cuaca dalam bentuk visualisasi dashboard yang informatif.

5. Manfaat Penelitian

Adapun manfaat yang diharapkan dari proyek ini adalah sebagai berikut:

5.1 Manfaat Akademis

- Menambah pemahaman mengenai konsep data engineering dan automated data pipeline.
- Menjadi bahan pembelajaran dan referensi dalam mata kuliah Rekayasa Perangkat Lunak.

5.2 Manfaat Praktis

- Memberikan contoh implementasi pengolahan data cuaca secara otomatis.
- Menyediakan dashboard cuaca yang dapat digunakan untuk analisis sederhana.

5.3 Manfaat Bagi Pengembang

- Meningkatkan keterampilan dalam penggunaan Airflow, dbt, Superset, dan Docker.
- Menjadi portfolio proyek data engineering end-to-end.

6. Batasan Masalah

Agar pembahasan dalam proyek ini lebih terarah, maka ditetapkan batasan masalah sebagai berikut:

1. Data yang digunakan hanya berasal dari Weatherstack API.
2. Sistem dijalankan pada lingkungan lokal menggunakan Docker.
3. Fokus proyek pada proses pengambilan, pengolahan, dan visualisasi data.
4. Tidak dilakukan prediksi atau analisis lanjutan berbasis machine learning.

7. Sistematika Penulisan

Sistematika penulisan laporan ini disusun sebagai berikut:

- **Bab I Pendahuluan:** Berisi latar belakang, rumusan masalah, tujuan, dan manfaat penelitian.
- **Bab II Metodologi:** Menjelaskan metode dan tahapan pengembangan sistem.
- **Bab III Implementasi dan Pembahasan:** Menjelaskan proses implementasi sistem.
- **Bab IV Penutup:** Berisi kesimpulan dan saran.

8. Penutup

Dengan adanya laporan pendahuluan ini, diharapkan pembaca dapat memahami latar belakang, permasalahan, serta tujuan dari pengembangan Automated Data Pipeline data cuaca. Laporan ini menjadi dasar untuk tahapan pengembangan dan pembahasan pada laporan selanjutnya.