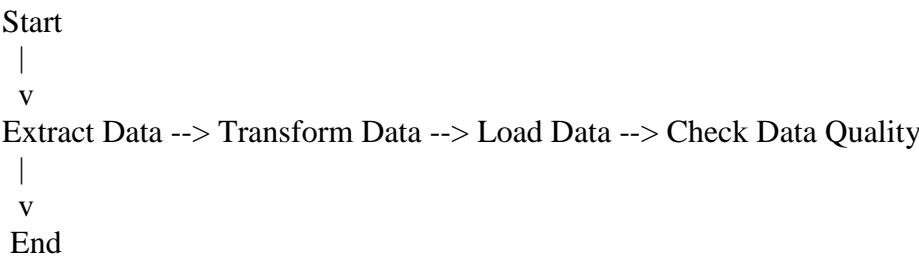


# Dokumentasi Implementasi Airflow: ETL Saham Brazil

## 1. Diagram Arsitektur DAG

Diagram arsitektur DAG menggambarkan alur proses ETL data saham Brazil, yang terdiri atas beberapa tugas utama:



Penjelasan Dependensi:

- a. Setiap task tergantung pada keberhasilan task sebelumnya.
- b. "Check Data Quality" dilakukan setelah data berhasil dimuat ke sistem.

## 2. Deskripsi Tujuan Setiap Tugas

Nama Tugas	Deskripsi
extract_data	Menarik data saham dari sumber eksternal (API atau file CSV).
transform_data	Membersihkan dan menormalkan data (misalnya: parsing tanggal, format angka).
load_data	Memuat data yang telah dibersihkan ke database atau data warehouse.
check_nulls	Memverifikasi apakah ada nilai kosong dalam kolom-kolom penting.

## 3. Informasi Penjadwalan dan Dependensi

DAG Name	Jadwal Eksekusi	Ketergantungan	Catchup
brazil_stock_market_etl	Harian (00:00)	Setiap task bergantung task sebelumnya	False
data_quality_brazil_stock	Harian (@daily)	Bergantung pada task check_nulls	False

- Catchup = False memastikan DAG hanya jalan saat aktif dan tidak memproses backlog.

#### 4. Pengaturan Pemantauan dan Peringatan

- *Web UI Monitoring:*
  - a. Setiap DAG dapat dimonitor melalui Airflow UI.
  - b. Status: Success / Failed / Running ditampilkan di visual DAG.
- *Health Check Scheduler:*
  - a. Diatur melalui konfigurasi airflow-scheduler, port 8974 untuk pengecekan kesehatan.
- *Logging:*
  - a. Log tiap task disimpan di level INFO atau ERROR, ditampilkan di tab "Log" per task.
- *Optional Alerts (jika diaktifkan):*
  - a. Notifikasi email atau Slack bisa dikonfigurasi melalui on\_failure\_callback.

#### 5. Prosedur Pemulihan Kegagalan

- a. *Task Retry:*

Semua task memiliki konfigurasi retries=1, artinya akan dicoba kembali sekali jika gagal.
- b. *Manual Trigger:*

Jika gagal, task bisa dijalankan ulang secara manual via Airflow UI.
- c. *Task Instance Clear:*

Admin bisa menggunakan fitur "Clear" untuk mengulang eksekusi task tertentu dari titik kegagalan.
- d. *Debugging:*

Log tersedia untuk setiap task guna menganalisis penyebab kegagalan.