

# Case Study ETL Airflow

Data Engineer

## Objectives

1. Melakukan extract dari data yang disediakan
2. Melakukan load database airflow
3. Membuat DAGs dan memilih operator yang sesuai

## Deskripsi

Create DAGs with Airflow

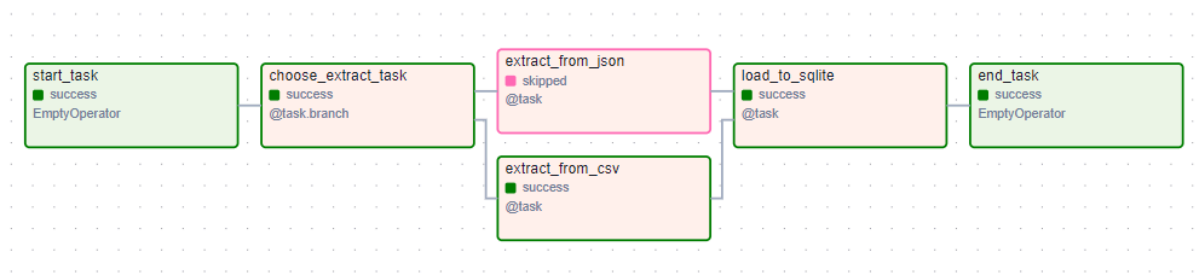
## Detail

**Aplikasikan EL sederhana pada Airflow dengan ketentuan berikut:**

1. Extract data dari berbagai sumber ke dalam sqlite
2. Menggunakan **Run Config Param** untuk menentukan extract data apa dan dari jenis source apa
3. Masing masing jenis file di-extract menggunakan task yang berbeda (gunakan **BranchOperator** untuk memilih ekstraksi menggunakan task yang mana)
4. Simpan data hasil ekstrak ke dalam file berjenis sama di staging area (**folder data/**). contoh: csv - semua data hasil ekstraksi akan disimpan sebagai csv tidak peduli sourcenya dari mana
5. load data dari staging area ke dalam sqlite menggunakan task yang berbeda
6. file dikumpulkan bersama dengan screenshot parameter + current graph + folder data + hasil di sqlite

### Tips:

1. Gunakan “**data/{namafile}.extention**” untuk mengakses file pada folder data
2. Gunakan db baru dengan cara **create\_engine(sqlite:///data/{namadb}.db)**
3. Format tipe data dict/json ke string
4. Contoh dag



## Tools

Visual Studio Code, Docker, Airflow