

Skill Task - Data Engineer

Anda akan diberikan 2 tasks yang perlu dikerjakan sebagai bagian proses data engineer

1. Diberikan data operasional dengan skema sebagai berikut - <https://dbdiagram.io/d/6069e955ecb54e10c33e9fea> (soal-1.sql). Buatlah data warehouse (DWH) dengan skema - <https://dbdiagram.io/d/5ff80a3380d742080a358222>. Anda diminta untuk membuat:
 1. DAG airflow (schedule satu hari sekali jam 7 pagi WIB) (pastikan jalan di airflow v2)
 2. File .sql dari DWH tersebut dan isi datanya dari tabel lampirannya
 3. (opsional) Skema DWH alternatif Anda -apabila ada- (dalam url dbdiagram.io) dan alasannya berdasar data operasional. Apa pendapat Anda dan masukkannya untuk DWH tersebut.

Diberikan dataset terlampir (soal-2.json). Lakukan cleansing, transformasi, dsb, lalu keluarkan list nama komoditas dengan sum berat masing-masing dalam satuan kg, terurut dari yang terbesar. Contoh keluaran program:

```
> ./run.sh soal-2.json
1. nila: 1500kg
2. lele: 1400kg
3. mas: 1300kg
4. kembung: 1290kg
...
```

2.

Keterangan

1. satu item pada list di soal bisa lebih dari satu komoditas, sehingga "nila, lele, mas" dan berat berisi "1kg, 2kg, 3kg", itu berarti nila 1kg, lele 2kg, dan mas 3kg.
2. buat programnya dalam bentuk aplikasi CLI. boleh dalam bahasa pemrograman apapun (python, javascript, go, etc). boleh juga pakai library apapun. pastikan ada:
 1. [README.md](#) dengan cara setup dan running program tersebut.
 2. Struktur kode dan modul
 3. History commit yang iteratif dan sesuai syntax <http://karma-runner.github.io/4.0/dev/git-commit-msg.html>
 4. Boleh gunakan asumsi pribadi selama bisa memberikan reasoning yang common sense. (cantumkan pada [README.md](#))

Waktu pengerjaan 3 hari dari email dikirimkan dan dikonfirmasi balik oleh Anda dengan membalas emailnya. Hasil tugas **disimpan di Github (satu repot) lalu kirimkan pranalanya** melalui email ke amalia.hardianti@efishery.com dan dimas.gs@efishery.com . Bila ada pertanyaan atau diskusi bisa membalas email ini.

Soal 1:

https://drive.google.com/file/d/1suAi4naRMZxW2vxzM_QGD_sYufgTyslh/view?usp=sharing

Soal 2:

<https://drive.google.com/file/d/1u0uvMs7NgKhwhkSNgASytQ18OYVZ0Fds/view?usp=sharing>