

# Data Science Challenge Gold Level

#### Hal-Hal yang Harus Kamu Lewati







Kriteria pengumpulan challenge yang harus kamu penuhi untuk dapat melewati level





# **Challenge:**

Membuat API untuk Cleansing Data dan Laporan Analisis Data



#### **SKILL METRICS**

(Kemampuan teknis yang akan kamu pelajari)

- Melakukan programming menggunakan python:
  - Operators
  - Data structure
  - Flow control
  - Function
  - File I/O
- Menggunakan SQLite dalam sebuah API dengan konsep RDBMS
- Menggunakan Pandas dan RegEx dalam melakukan cleansing data
- Menggunakan matplotlib dan seaborn dalam melakukan visualisasi data
- Menggunakan proses berpikir analisis data untuk mencapai objektif bisnis dalam membangun API
- Menggunakan Flask dan Swagger UI dalam membangun API
- Menggunakan Github dan Git
- Menggunakan power point atau slides tool untuk membuat laporan



#### **DELIVERY**

(Hal yang akan kamu lakukan untuk dapat melewati level)

- 1. Membaca dan memahami studi kasus yang diberikan.
- 2. Menggunakan fitur-fitur python dan Pandas serta RegEx untuk membersihkan data
- 3. Memasukkan hasil koding (source code) cleansing data ke Flask
- 4. Menggunakan Flask dan Swagger UI untuk membuat API
- 5. Menggunakan SQLite sebagai database di dalam API
- 6. Membangun API yang mampu mengakomodir input user berupa teks dan file
- 7. Membangun API yang mampu mengakomodir pemrosesan text cleansing dengan output berupa teks yang sudah di-cleansing
- 8. Menggunakan Git dan Github untuk kolaborasi
- 9. Membuat laporan (report) analisis data dari teks



#### CRITERIA

(Kriteria pengumpulan challenge yang harus kamu penuhi untuk dapat melewati level)

- 1. Terdapat API yang bisa menerima input user berupa teks dan file dengan 2 endpoint (10%)
- 2. Server API dibuat dengan Flask dan Swagger UI (25%)
- 3. Penyimpanan data dalam SQLite menggunakan modul SQLite 3 (15%)
- 4. API yang bisa menghasilkan output berupa teks yang sudah di-cleansing (10%)
- 5. API yang bisa memproses text cleansing (15%)
- 6. Membuat report (laporan) dari hasil analisis data teks (25%)



# **API untuk melakukan Data Cleansing**

Kamu baru aja direkrut jadi Junior Data Scientist. Nah, usermu membutuhkan API untuk membersihkan data teks secara otomatis.

Teks yang bakal dibersihkan pun berupa teks non formal.

Terus langkah apa yang harus dilakukan?





# Masuk ke langkah pertama dalam pengerjaan API~

Pertama, kamu perlu buat **endpoint dari API** yang isinya adalah sebagai berikut:

- 1. Input teks via form, lalu outputnya berupa teks yang sudah dibersihkan
- 2. Upload file CSV, lalu outputnya berupa teks yang sudah dibersihkan





Lalu kamu ambil data teksnya di <u>link</u> ini. Setelah itu pastikan sudah instal library Flask, Swagger UI, dan Pandas, dan RegEx.

Kamu harus membuat endpoint yang bisa menerima input user berupa text dan file. Lalu kamu buat fungsi untuk melakukan data cleansing.





#### Hasil akhir project ini ada dua:

- API yang dapat memproses text cleansing dan menghasilkan output berupa teks yang sudah di-cleansing.
- 2. **Laporan hasil analisis data** dari data teks tersebut.





Untuk laporan (report) mengandung bagian seperti di bawah dengan 1 slide mengandung 1 bagian:

- Judul dan nama ((1 slide)
- Pendahuluan (1 slide)
- Metode Penelitian (1 slide)
- Hasil dan Kesimpulan (1 slide)





Kalau sudah selesai dengan project API yang kamu kerjakan maka kamu **upload ke Github** dengan akun pribadimu.





### Saatnya mengumpulkan project~

Oke, setelah upload ke Github sudah selesai kamu masih harus mengumpulkan file project tersebut sob, dengan langkah sebagai berikut:

- Copy paste link Github tempat kamu mengupload project dan masukkan ke <u>Google</u> Form
- 2. Kumpulkan file project (API dan report) ke dalam bentuk .zip
- 3. Upload .zip tersebut ke Google Form yang sama tadi sob.





Setelah API dan laporan kamu lengkap, siap-siap ya buat presentasi di kelas biar user paham tentang hasil analisis datanya.

Setelah itu, hasil kerjamu akan dinilai oleh fasilitator sebagai challenge Level Gold di Data Science Course ini.

Selamat belajar dan good luck! 😉





Terima kasih

