

## [CSV Query Language] File Processing and Search Study Case

---

Dalam setiap perusahaan yang memiliki data baik itu berupa data nasabah, pegawai maupun data2 lainnya akan disimpan dalam sebuah format tertentu berupa sebuah file atau bahkan dengan sistem penyimpanan data yang lebih kompleks (misal database system).

Sebagai contoh sebuah perusahaan dibidang jasa finansial menyimpan data form aplikasi peminjaman uang ke dalam sebuah file dengan format CSV (comma separated values). File ini bisa berisi ribuan lebih data pengajuan peminjaman pelanggan dan untuk mencari data dalam file tersebut akan memakan waktu yang cukup lama jika dilakukan secara manual, maka dari itu anda diminta membuat sebuah solusi program pencari data menggunakan sintaks pencari sendiri (Query Language) untuk file berformat CSV. Query language adalah sebuah bahasa khusus dalam dunia teknologi komputer yang digunakan untuk mencari sebuah data dalam suatu sistem. Contoh query language yang paling populer adalah SQL (database system), GraphQL (Facebook API), XPath (XML document), dan sebagainya (see [this link](#) for details).

Berikut contoh data yang akan kita pakai. Data tersebut jika dibuka menggunakan Microsoft Excel akan berbentuk seperti table berikut:

id	first_name	city	gender	has_car	has_property
1001	Brian	JKT	M	N	Y
1002	Lisa	BDG	F	Y	Y

\*id : nomor unik untuk setiap row data / per individu (seperti no ktp)

\*has\_car dan has\_property direpresentasikan dengan boolean value (Y = Yes, N = No)

### [Task 2A]

---

Buatlah sebuah program dalam C untuk mengambil input sebuah bahasa query dengan format berikut:

```
SHOW <coloum-filter> WHERE <search-key> equals <search-value>
```

dan jika dieksekusi akan mencari data dari file csv tersebut dengan kriteria/filter dari klausa **WHERE**. Berikut contoh hasil pencarian data:

**Query:** SHOW first\_name,city WHERE gender equals F

**Result:**

first\_name: Lisa, city: BDG

**Query:** SHOW \* WHERE id equals 1001

**Result:**

first\_name: Brian, city: BDG, gender: M, has\_car: N, has\_property: Y

note: \* digunakan untuk menampilkan semua kolom

**Query:** SHOW id,first\_name,city

**Result:**

id: 1001, first\_name: Brian, city: JKT

id: 1002, first\_name: Lisa, city: BDG

note: tanpa klausa **WHERE** maka akan menampilkan semua row data

### [Task 2B]

---

Dalam sistem yang memproses sebuah data yang besar, performance adalah hal yang paling diperhitungkan. Salah satu proses yang paling mempengaruhi performance pada program ini adalah setiap melakukan pencarian data, kita akan melakukan iterasi membaca file mulai dari row pertama hingga terakhir, atau dengan kata lain kompleksitas algoritmanya adalah  $O(n)$ . Coba **Analisa** dan berikan **Solusi** bagaimana cara meningkatkan metode pencarian data dalam sebuah file tersebut agar pencarian bisa lebih efektif, semisal mencari data dengan id=1002 hanya akan membutuhkan 1 kali pencarian /  $O(1)$ . Implementasikan dalam code jika memungkinkan.

Notes:

---

- File data yang dipakai (customer\_app\_data.csv) akan diupload di forum secara terpisah
- Tuliskan pertanyaan anda di forum jika membutuhkan penjelasan lebih detail
- Pengiriman jawaban final (code) kepada Dosen terkait, bukan di dalam forum
- Jika memposting code snippet (penggalan saja) masih diperbolehkan di dalam forum
- Diskusi terkait solusi **Task 2B** terbuka di dalam forum.