

**Guião I**  
**Laboratórios de Informática III**  
**2020/2021**

**Objetivos:**

- Manipular grandes volumes de dados em C;
- Consolidação de conhecimentos essenciais da linguagem (leitura e parsing de dados, I/O, gestão de memória);

**Ficheiros necessários:**

- users.csv (*user\_id*)
- reviews.csv (*review\_id*, *user\_id*, *business\_id*, *stars*, *useful*, *funny*, *cool*)
- business.csv (*business\_id*, *name*, *city*, *stars*, *review\_count*)

**Exercícios:**

1. Dados os ficheiros de dados disponíveis na BB, faça um programa na linguagem C que permita ler cada um dos ficheiros e guardar o seu conteúdo em memória, usando o tipo de dados char \*\*. Após a leitura e carregamento dos ficheiros, o programa deve conseguir escrever cada um dos registos contidos nos ficheiros no *stdout*, linha a linha.
2. Repita o procedimento do exercício anterior, mas guardando agora cada registo dos ficheiros num tipo de dados mais apropriado (struct user, struct review, struct business), que contém os campos contidos em cada registo.
3. Reutilizando o código do exercício anterior, efetue a validação da informação inferida dos registos dos ficheiros e apresente as seguintes estatísticas relativas ao seu conteúdo:
  - a. Quantos negócios (business) têm reviews?
  - b. Quantos negócios têm 4 ou + estrelas?
  - c. Qual o negócio com mais *reviews*?
  - d. Qual a cidade com maior número de negócios?
  - e. Quais as 3 *reviews* mais “cool”?
  - f. Quantos *users* fizeram reviews?
  - g. Quanto *users* não fizeram qualquer *review*?
  - h. Quantas *reviews* inválidas existem (*reviews* cujos id 's não existam nos registos de business e users )?