



Considere os ficheiros *clientes.csv* e *vendas.csv*.

1. Faça a leitura dos ficheiros.

2. Detete e corrija os seguintes problemas:

- Duplicados (linhas repetidas).
- Nulos (campos vazios).
- Idades inválidas (idade < 18 ou ausente).
- Emails mal formatados (sem @).

Apoio:

- Ler o ficheiro CSV com pandas.
- Detetar e remover duplicados (com base no nome e email).
- Tratar valores nulos ou vazios no campo email (substituir por "email_desconhecido").
- Corrigir idades inválidas (idade < 0 ou > 120 → substituir por None).

3. Valide a integridade referencial: apenas deve considerar clientes que têm *id_cliente* usado em vendas.

4. Guardar o resultado em *clientes_final.csv*.

5. Faça agora a leitura dos ficheiros *clientes_final.csv* e *vendas.csv*.

6. Valide a integridade referencial: apenas manter vendas cujo *id_cliente* exista na tabela de clientes.

7. Calcule o total de vendas válidas.

8. Guarde o ficheiro final *vendas_validas.csv*.