

Recorrendo apenas às cláusulas anteriores não é possível obter tabelas com valores estatísticos (somatórios, médias, etc.). Tal acontece porque as operações de agregação (*sum*, *avg*, *count*, etc.) não podem ser aplicadas linha a linha, apenas sobre um grupo de linhas.

Por exemplo, o comando para listar os códigos postais onde há mais que dois clientes não pode ser efectuado da seguinte forma:



```
SELECT CodPostal FROM Cliente WHERE COUNT(CodPostal) > 2
```

O comando é incorrecto porque a cláusula WHERE é testada linha a linha e numa só linha não é possível obter o total de códigos postais.

As cláusulas GROUP BY e HAVING permitem manipular valores agregados.

- A cláusula GROUP BY permite segmentar os registos de uma tabela em vários grupos e fazer um cálculo agregado para cada grupo;
- A cláusula HAVING é equivalente à cláusula *where*, mas as expressões lógicas com que trabalha são verificadas para cada grupo criado pela cláusula GROUP BY; tipicamente, estas condições avaliam o resultado de valores agregados ( $\text{count}(* > X$ ,  $\text{sum}(\text{valor}) < Y$ , p. ex.).

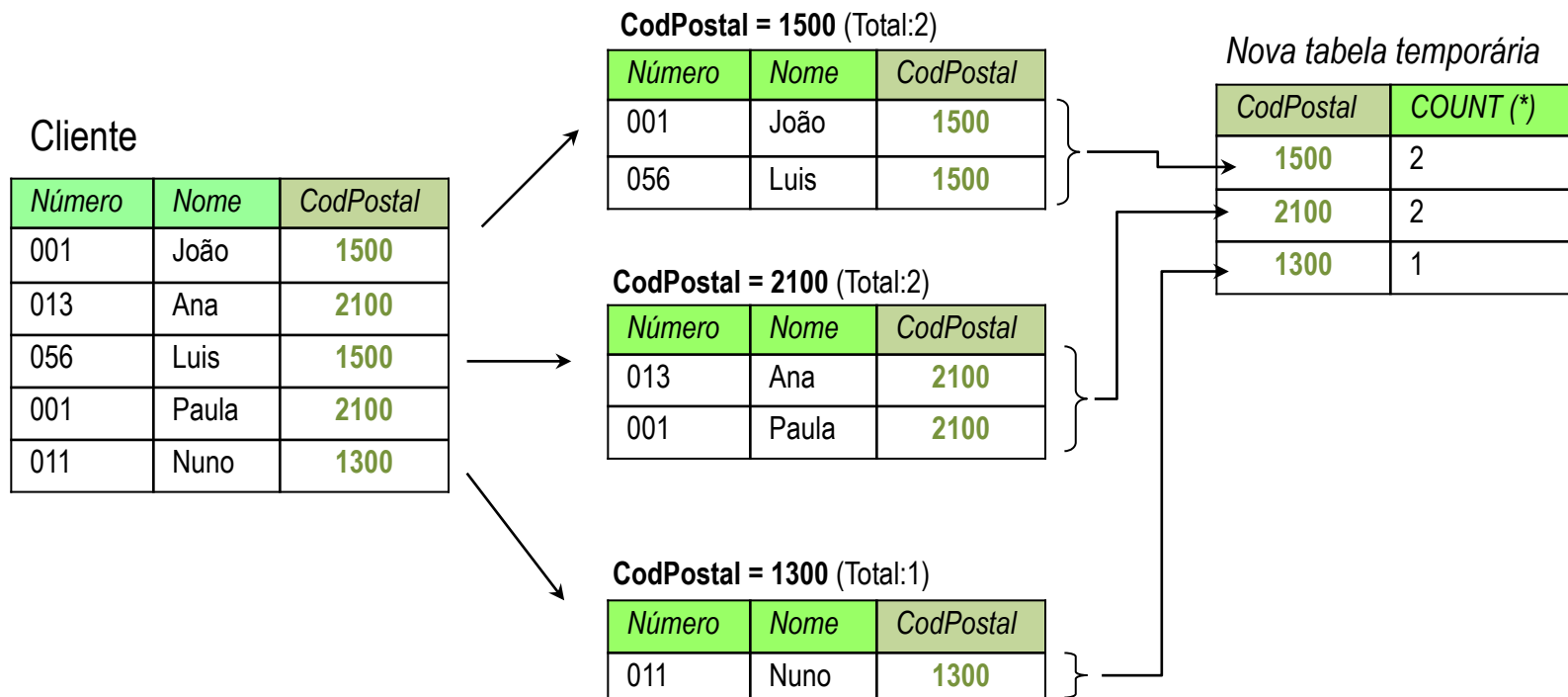
O exemplo anterior será bem formulado da seguinte forma:

```
SELECT CodPostal FROM Cliente
  GROUP BY CodPostal
  HAVING COUNT(*) > 2
```

- A cláusula GROUP BY segmenta os clientes por código postal, produzindo tantos grupos de clientes quantos os códigos postais distintos onde existem clientes;
- A cláusula HAVING faz uma filtragem para que apenas se seleccionem os grupos com mais do que 2 clientes.

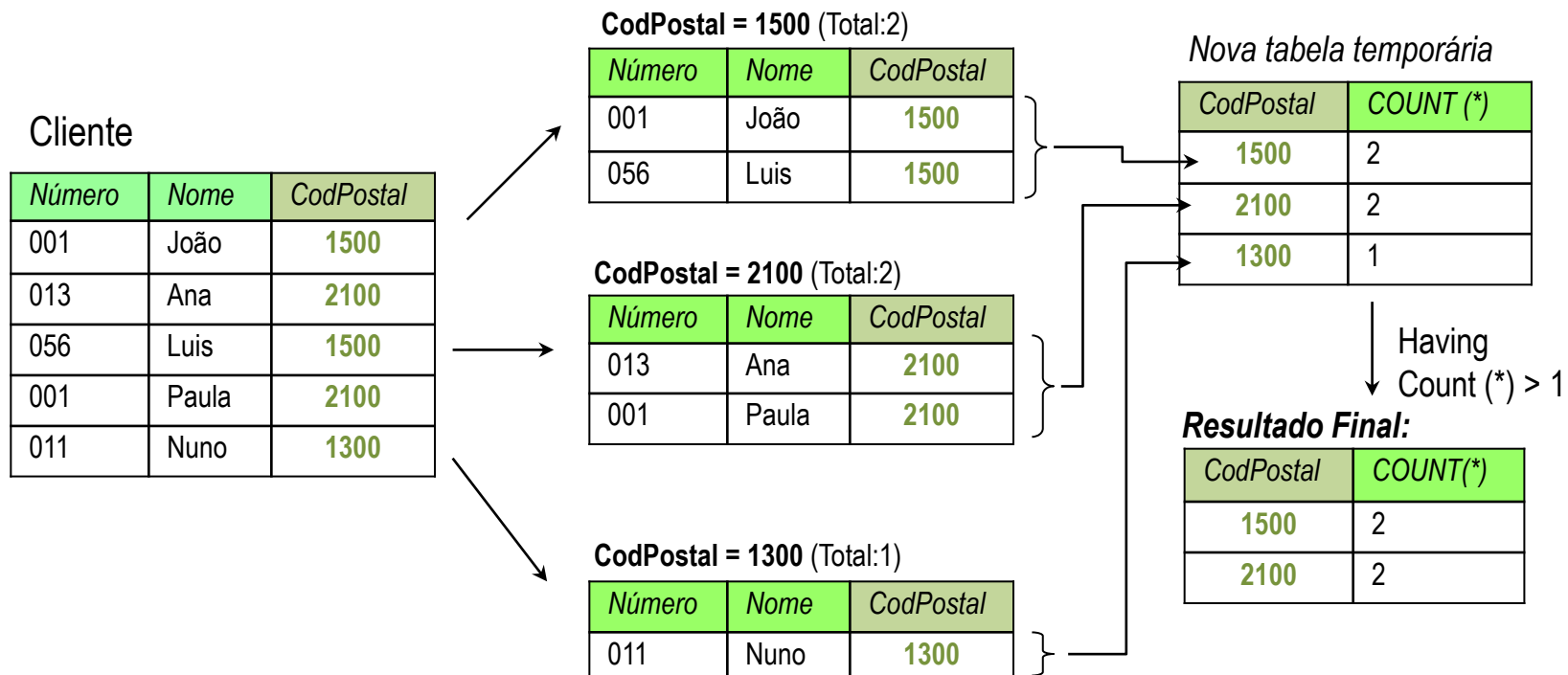
O seguinte comando lista, para cada código postal, o número de clientes que a ele estão associados, desde que exista mais do que um cliente:

```
SELECT CodPostal, COUNT(*)
FROM Cliente
GROUP BY CodPostal
```



O seguinte comando lista, para cada código postal, o número de clientes que a ele estão associados, desde que exista mais do que um cliente:

```
SELECT CodPostal, COUNT(*)
FROM Cliente
GROUP BY CodPostal HAVING COUNT(*) > 1
```





Sempre que se coloca GROUP BY numa instrução, na cláusula *select* desta apenas podem constar expressões agregadas (*count*, *sum*, etc.) e as colunas que estão na cláusula GROUP BY.