

04

## Problema comum

Um problema muito comum que encontramos ao tentar juntar dois bancos de dados num só é não haver uma variável comum entre eles. Mais precisamente, podemos até ter os mesmos dados em duas variáveis em bancos diferentes, mas elas não têm o mesmo nome. Na aula, nos deparamos com esse problema.

The screenshot shows the RStudio interface with the console tab active. The console window displays a series of '>' symbols followed by a merge command:

```
> popularidade_e_duracao <- merge(sumario_estatistico, popularidade, by = "curso")
Error in fix(by(by.x, x)) : 'by' must specify a uniquely valid column
>
>
```

A red rectangular box highlights the error message: "Error in fix(by(by.x, x)) : 'by' must specify a uniquely valid column".

Temos o banco `sumario_estatistico` com as variáveis `Group.1`, que traz o código dos cursos, e `x`, que traz o número de dias que os cursos levam para serem concluídos em média, e o banco `popularidade` com as variáveis `curso` e `popularidade`. Para resolver esse problema da indexação, precisamos escolher um dos bancos de dados e alterar o nome da variável indexadora. Como podemos fazer isso?

Selezione 2 alternativas

**A** `sumario_estatistico <- rename(sumario_estatistico, c("Group.1"="curso","x"="dias"))`

**B** `popularidade <- rename(popularidade, replace = c("curso"="Group.1", "popularidade" = "x"))`

**C** `sumario_estatistico <- rename(sumario_estatistico, c("Group.1"="dias","x"="curso"))`