

Relacionando tabelas parte 2

Transcrição

[00:00] Continuando a parte de relacionamento entre diferentes tipos de tabelas, eu vou mostrar agora como é que... qual que é o processo para a gente fazer outros tipos de relacionamentos. A gente já mostrou como é que funcionava o left join, era o primeiro, agora, a gente vai fazer outros tipos de relacionamentos.

[00:20] Lembrando, a tabela que nós estamos trabalhando aqui dentro do R, essa tabela é de cargo e local, estou fazendo relacionamento entre essas duas tabelas, a primeira tabela onde eu tenho o nome dos funcionários, o cargo dos funcionários e uma segunda tabela, eu tenho o nome dos funcionários e o local onde eles trabalham.

[00:38] Então, agora, vamos trabalhar aqui com o right join, a gente vai fazer esse relacionamento pela direita, pela tabela da direita, que é a tabela do local, eu vou manter essa tabela local e vou buscar na tabela da esquerda, o respectivo representante.

[00:55] Então, a hora que... depois que a gente fizer esse relacionamento, usando o right join, vai acontecer o seguinte, eu trago toda a minha tabela local, então eu tenho aqui Fabio, comercial, Marcos do CRM, a Fernanda de TI e a Julia do RH. E aí, eu vou olhar o cargo que esse funcionário ocupa.

[01:16] Eu não encontrei o funcionário Fabio, porque ele não está na tabela de cargos, os demais, eu consigo encontrar. Então, vamos fazer isso no R, vamos ver como é que fica, vamos replicar exatamente esse relacionamento dentro do RStudio. Então, agora eu tenho o right join, vou abrir um chunk aqui para a gente fazer a análise.

[01:47] Então, como é que vai fazer? Eu já tenho as minhas tabelas, cargo e local criadas aqui dentro do R, a gente criou exatamente... criando, imputando essas listas, então já está criado. Então, agora é só fazer o quê? Eu vou lá e... cargo, eu vou fazer... vou trabalhar com o pipe, então, vou adicionar aqui o "Alt"... Opa, pulou.

[02:20] Vou adicionar uma... deixa eu só... isso... vou colocar um (centro) aqui para ficar melhor a visualização, "Ctrl + Shift + M". E aí, agora, eu vou colocar exatamente o right join, relacionando agora pela minha tabela local e adicionando a chave, a chave que eu estou querendo cruzar aqui, que é o by nome, o campo comum aqui entre as duas tabelas, que é a chave que a gente vai cruzar isso aqui.

[02:56] Então, mantive a primeira tabela, mantive o nome, Fabio, Marcos, Fernanda e Julia, não encontrei o cargo do Fabio, mas eu sei o local da onde ele vem, que é comercial. O Marcos é analista, trabalha no CRM, Fernanda, coordenadora em TI e Julia... Então, é exatamente isso que eu não encontrei, não encontrei o cargo do Fabio.

[03:26] Eu sei que o local, eu trouxe, local é comercial. Então, aqui o right join, ele faz o relacionamento mantendo como referência a minha tabela da direita, minha segunda tabela. Agora, um outro tipo de relacionamento é chamado... a gente vai ver aqui o inner join.

[03:51] O inner, é como se nós tivesse fazendo a intersecção desses... de novo, a mesma tabela, (aliás), da mesma tabela que a gente vai criar, quando a gente criar o inner, o resultado vai ser o quê? Vai ser exatamente a intersecção dessas tabelas, ou seja, o que é comum nessas tabelas.

[04:08] Ele elimina aqueles valores ausentes, valores ausentes, agora não vai aparecer, você vê que não apareceu nenhum ausente aqui, ele trouxe só o que eu encontrei nas duas tabelas, seria como se fosse o comum. Então, vamos duplicar esse comando.

[04:25] Então, de novo, tenho a minha tabela cargo, “Ctrl + Shift + M”. Opa, aqui não, “Ctrl + Z”, eu coloco aqui um pipe, para passar para a linha de baixo o comando, só lembrando isso. Então, usando o inner join, o inner join pela tabela local e coloco a minha chave, igual, abre, fecha aspas, nome e vou executar aqui.

[04:56] Então, agora eu estou aplicando o comando inner join. Aplicando o inner join, traz o que é comum, veja que agora não tem nenhum valor ausente nessa relação, nesse relacionamento, ele pega, excluiu o Fabio do comercial, por que que ele excluiu o Fabio?

[05:15] Porque a hora que eu cruzo essas duas tabelas, eu não consigo ter esse valor comum nas duas tabelas. Então, o inner join, ele traz somente o que é comum. Próximo... Existe um outro tipo de relacionamento agora, que é gull join. O full join, ele vai trazer as duas tabelas que você está relacionando.

[05:38] Então, como o nome diz, vai trazer todas as informações da tabela “A” e todas as informações da tabela “B”, todas as informações da minha tabela cargo e todas as informações da tabela local. O resultado desse relacionamento, usando essa função full join, vai ser o seguinte, então ele vai trazer aqui, está vendo?

[05:56] Ele vai trazer o André, analista, ele não traz nada no local, o Marcos, ele consegue achar, Fernanda, coordenadora, acha no TI, a Julia, acha o Fabio, não consegue achar o cargo, mas traz o comercial, ou seja, então ele relacionou essas duas tabelas, tanto da esquerda, quanto da direita.

[06:15] Vamos ver como é que fica lá no RStudio, fazer esse comando. Então, agora, vamos usar aqui o comando... vou escrever aqui, a gente vai estar utilizando agora o join. Mesma coisa, eu vou ter aqui o cargo ou... de novo, vou ter o cargo, agora a gente vai colocar aqui o comando full join, mesmo esquema, relacionando a tabela local, pela chave by, igual a nome.

[07:10] E aí, vou executar. Executando, veja, replicou exatamente a tabela que a gente mostrou aqui no PPT, ele traz a junção das duas tabelas, deixando ausente aqui, o cargo ele não encontra, porque não tem o cargo do Fabio, do funcionário Fabio e também, ele deixa o valor ausente no local, aqui, porque ele não tem o local que o funcionário André trabalha.

[07:40] Agora, tem um outro... um último relacionamento, que um dos mais... é um pouco menos usual. Esses três... o right join, o left join, o inner join e full join, são os mais comuns, a gente utiliza quando a gente está relacionando tabela. Tem um outro relacionamento chamado anti join, que a gente usa, que a gente... que dá para utilizar aqui no R, ele vai justamente fazer o quê?

[08:12] Ele vai separar aquilo que você não conseguiu juntar, isso pode ser feito de duas maneiras, dependendo da chave que você está utilizando. Então, a gente vai pegar e abrir aqui, vou mostrar para vocês como é que faz o anti join. O Anti join, mantendo as mesmas duas tabelas, eu vou usar o comando anti join e eu quero trazer aquilo que eu não encontro na relação.

[08:38] Então, deixar somente como resultado, aquilo que eu não encontrei, é isso que ele vai fazer. Então, primeiro, trazendo aqui, relacionando, mantendo a tabela como referência, cargo e eu vou na local e trago e falo: o que que eu não encontrei da tabela cargo, na tabela “A”, que eu não encontrei na tabela “B”, que é local.

[09:01] Ou seja, eu não encontrei o André analista, o André com cargo de analista, porque ele não está nessa segunda tabela. Ele seleciona, ele separa para você, exatamente aquilo que você não está encontrando na relação. Vamos fazer esse comando no RStudio, para ver como é que vai ficar.

[09:21] Então, aqui, cargo e aí, eu vou fazer esse comando aqui do anti join. Então, anti join pelo local, usando a chave... mesma chave que a gente está utilizando aí, nome. Então, executar, veja bem, então trouxe exatamente só uma observação, que é uma linha, André com cargo de analista.

[10:04] Agora, se a gente quisesse pegar e fazer ao contrário, se eu quisesse trazer tudo, fazer essa relação invertida agora, eu vou olhar na minha tabela local e ver tudo aquilo que eu não consigo encontrar na minha tabela cargo, seria ao contrário. Então, veja bem, aqui a gente trocou.

[10:23] Olha a diferença, aqui eu estou tomando como referência a minha tabela local e faço um anti join na tabela cargo, o que que ele retorna? Ele vai retornar o funcionário Fabio do comercial, por quê? Eu não tenho essa relação do outro lado, então ele está trazendo, selecionando o que eu não estou encontrando, o anti join, ele traz aquilo que você não encontrou e deixa como resultado.

[10:46] Então, vamos fazer esse segundo relacionamento, eu vou só complementar aqui então, esse primeiro é o anti join pela primeira tabela, que a gente está (utilizando), então seria cargo e agora, a gente vai fazer... vou copiar aqui, colar, para ficar mais fácil, mas vou fazer o que, agora?

[11:15] Vou fazer pela segunda tabela, que agora é local e o que que eu mudo? Eu vou inverter esse campo, agora na minha tabela referência local e eu estou buscando na minha tabela cargo. Então, isso vai me devolver o quê? Vai me devolver somente o Fabio do comercial, pois eu não consigo achar ela na minha tabela cargo.

[11:45] Executei e aí, devolve exatamente Fabio comercial. Então, legal, esses tipos de relacionamentos são bem comuns quando nós estamos trabalhando com cruzamento de diferentes tabelas. Aqui, eu mostrei apenas cruzamento por uma chave comum, mas eu poderia usar outras chaves.

[12:07] Então, tabelas mais complexas, onde você tem, por exemplo, o CPF da pessoa e aí, você cruza o CPF com uma data ou o CPF com o mês, então, geralmente você pode usar essa chave dupla também, CPF, RG, data, chave tripla.

[12:25] Então, você pode usar “n” chaves aqui para cruzar isso aqui em diferentes tabelas, mas o importante é vocês entenderem exatamente esses tipos de relacionamentos, quando a gente vai fazer essa relação matemática, entre diferentes tipos de tabelas, o left join, o right join, o inner join, o full join e o anti join, também pode ser bastante útil para separar aquilo que você não está encontrando exatamente nesse relacionamento de tabelas.

[12:59] Então, era isso aí.