

06

Em busca da chave única

Transcrição

[00:00] Depois desses resultados que a gente mostrou para o nosso cliente, ele passou uma informação nova que conflita com o que falamos para ele. Quando fizemos nossa frequência de jogos no categoria, tinha dado 41 jogos diferentes. Mas o cliente me disse que esse número não bate com o que ele tinha no catálogo dele. Ele disse que tinha 42 jogos diferentes.

[01:47] Se a variável nome não está me dizendo realmente quantos jogos diferentes temos, isso significa que existe outra característica do jogo que diferencia um do outro que não só o nome. Existem jogos diferentes com nomes iguais.

[02:34] Usando meu PROCFREQ, vou fazer uma frequência cruzada entre meu nome e a data de lançamento ou nome e gênero. Vou usar gênero por enquanto porque são menos categorias.

[02:57] Agora quero fazer uma nova tabela de frequência cruzada. Preciso pedir para o PROCFREQ escrever do zero? Não, posso passar o comando table duas vezes no mesmo FREQ, porque quero partir da mesma base. Executando, vão aparecer duas tabelas ao mesmo tempo.

[04:02] Não parece ser o jeito mais simples de observar. Existe uma opção do SAS que facilita? Não quero ver em forma de matriz. Poderia ser em forma de lista e eu veria o jogo que aparece duas vezes.

[04:55] Voltando para o código, usando as opções da tabela de frequência, coloco a barra e escrevo list, que é lista em inglês. Executando novamente, minha última tabela tenho dois jogos com o mesmo nome com gêneros diferentes. Já sei meu número real de jogos.

[06:33] Descobrimos que a variável nome não é suficiente para eu saber quantos jogos diferentes eu tenho. Se eu quiser saber, preciso usar pelo menos duas variáveis. Dizendo em termos mais técnicos, minha variável nome não é uma chave única. Preciso usar duas, como nome e gênero. Essas duas variáveis juntas formam uma chave única, em que se eu usar, não vou ter duplicações de jogos, como acontecia antes.