

PERGUNTAS REALIZADAS DURANTE A AULA E RESPONDIDAS POR E-MAIL

Curso: MBA em Data Science e Analytics

Disciplina: Data Wrangling II

Data: 16/11/2021

Victor Bastos da Silva

nesse comando da função `right_join` faltou ele descrever o dataset y, não?
`right_join(base_covid, by = "nome") %>%`

Resposta: Bom dia! No uso do pipe, o primeiro argumento fica fora da função. Neste caso, o primeiro argumento é o dataset X na função `right_join`. Então, no comando `- dataset_inicial %>% right_join(dataset_merge, by = "observacoes")` – o dataset_inicial é o X (está fora) e o dataset_merge é o Y.

Michel Faleiros Martins

Mas, a nota 100 não deveria aparece no `ajuste_rotten_3`?

Resposta: Bom dia! Está correto, o comando foi ajustado e disponibilizado no material atualizado.

Anderson Jeronimo Da Silva

Podemos agregar a `gsub` à `str_sub` através de Pipe?

Resposta: Neste caso, a melhor forma de integrar as duas funções é colocar uma como argumento da outra. Por exemplo, no caso do exercício 6, poderíamos fazer:
`ajuste_rotten <- gsub("/", "", str_sub(completo$Rotten.Tomatoes, 1, 3))`
Então, a `str_sub` entra como argumento na `gsub`.

Vanessa Cristina Natel Sales

Não entendi qual a principal diferença do `filter` com e sem a função

Resposta: Bom dia! São apenas formas diferentes de executar o procedimento de seleção de linhas com base em critérios. Iniciamos verificando como pode ser feito utilizando o R Base e, na sequência, utilizamos a função `filter` do tidyverse. Creio que a função `filter` tende a ter uma aplicação mais simples.

Diego Ferreira Mello

Professor, por que não converter para o mesmo padrão nas duas variáveis, ex: 7.8 -> 78/100 ou o contrário ? e como fazer isso ?

Resposta: Bom dia! Neste caso, foi melhor extrair apenas os números. Isto ocorre, pois o objetivo é transformar a variável em números para que possamos calcular as estatísticas das notas. O padrão 78/100 está como texto, impossibilitando o cálculo.

Aquila José Gonçalves Delfino

Eu tenho uma duvida da aula passada, por que no comando `select`, o `everything()` desconsidera as variáveis que já foram solicitadas? Não deveria o `everything()` trazer todas as variáveis?

Resposta: Bom dia! Neste caso, o comando “entende” que as variáveis que foram especificadas por seus nomes ficam fora do comando everything(). Por isto, não repete as variáveis já especificadas por nomes.

Millena Santos

Com relação a junção das funções group_by e filter: Estando na ordem "group_by - filter" o programa filtra dentro dos grupos. Na ordem inversa: "filter - group_by" o programa agrupa depois de filtrar?

Resposta: Bom dia! Sim, exatamente isto. Quando utiliza o pipe, a ordem das funções altera o resultado. Neste caso, se pedir primeiramente o filter, os grupos serão formados no banco de dados já filtrado.

Rogério De Assiz

O "bind_rows" é posicional, ou seja, obedece a posição da variável ou considera o nome da variável no momento de unir os databases?

Resposta: Bom dia! É feito com base na posição das colunas. Isto significa que é necessário que as colunas (variáveis, por exemplo) estejam na mesma posição e sequência nos dois datasets.

Hernandes Matias Junior 124799-11640