

Organizando e Criando Novos Campos

Transcrição

[00:00] Vamos continuar o nosso projeto de análise de dados, utilizando a base que a gente criou com a seleção e os filtros de procura de imóveis, que a área de planejamento passou para a gente com alguns filtros, que a gente já fez. A base que a gente vai estar continuando a análise aqui é essa “casas_f2”, só relembrando.

[00:27] Então, na aula de hoje, nós vamos aprender como organizar a base de dados e fazer... criar alguns outros campos nessa base, só relembrando. Então, a área de planejamento passou alguns requisitos para a gente, para a gente achar os imóveis disponíveis de aluguel.

[00:48] E agora, nós vamos fazer uma organização na base de dados, criar alguns novos campos, para devolver isso para a área demandante de forma mais clara e objetiva. O primeiro passo aqui, só lembrando, eu vou sempre apontar a (lib name) que a gente está usando aqui, que é a tidyverse.

[01:17] Então, eu vou executar aqui ele, já deixar executada. A gente vai, primeiro passo aqui, a gente vai ordenar a base de dados, quando a gente ordenada a base de dados, a gente tem que declarar essa ordenação por alguma variável dominante e sempre aqui no R, ele vai dar como default, essa ordenação de modo crescente.

[01:52] O comando para a gente fazer isso é o arrange, vamos utilizar aqui como a casas_f2, que são as nossas casas, que já fizemos os filtros selecionados, requisitos... número de quartos, vagas, que área a demandante passou para a gente e aqui eu vou ordenar então pelo valor, os valores disponíveis.

[02:16] Então, eu dou arrange, casas_f2 e valor. Eu vou executar. Como a gente pode ver, a base foi ordenada, só lembrando, essa base casas_f2, a gente... nós tínhamos encontrado aqui 21 imóveis disponíveis para alugar. Então, ele aparece aqui na tabela ordenando do menor para o maior, sempre do menor para o maior, lembrando que o valor também era sempre abaixo de quinze mil.

[02:47] Como eu falei, então o R, ele sempre ordena de maneira crescente, se tivermos interessados em fazer essa disposição em maneira decrescente, então vamos colocar aqui ordenação de maneira... ordenação decrescente. Então, como é que a gente faz?

[03:13] Pega exatamente o mesmo comando que a gente digitou aqui em cima, arrange, casas_f2, só que o que que a gente vai fazer agora? É só colocar um desc do valor de maneira decrescente. Vou executar essa linha aqui com “Ctrl + Enter”, executei.

[03:33] Então, ele ordenou aqui de quinze mil para o menor, de quinze mil, até o menor valor aqui que é cinco mil e novecentos. Então, essa é a maneira de a gente organizar a base de dados, ordenando geralmente por uma chave, por um campo e a gente diz aí se a gente quer de maneira crescente ou decrescente.

[03:54] Uma outra coisa que a gente vai aprender agora é criar novos campos na base, então a gente já tem a base aqui, vamos criar novos campos, o... Então, eu vou colocar criando novos campos na base dados. O comando que a gente cria é o mutate. Então, o comando que a gente vai estar utilizando aqui para fazer isso, ele chama mutate.

[04:26] Então, vamos fazer o seguinte, a área de planejamento de dados também, ela passou um indicador para nós, um indicador sobre área e valor, de área sobre valor e ela quer fazer as visitas quando esse indicador for maior que cinco, por exemplo.

[04:48] Então, ela tem o número chave dela e ela quer verificar, quando esse indicador ultrapassar cinco, ela visita, senão, não visita. Então, o nosso objetivo aqui é criar esse indicador com esses parâmetros e depois a gente devolve a base para a área demandante com esse indicador.

[05:03] Então, vamos criar esse indicador da seguinte forma, eu vou salvar aqui na mesma base, casas_f2, ela vai receber... Então, o que que eu vou fazer? É o comando `mutate`, que a gente vai utilizar para a gente criar... O comando `mutate`, a gente cria novos campos ou modifica campos na base de dados.

[05:29] Então, a gente vai criar o indicador... Então, `mutate`. Ah, de novo, primeiro eu tenho que passar aqui aonde eu estou criando, aonde eu vou modificar ou criar. Então, `casas_f2` mesmo, na mesma base, criando o campo chamado de indicador. Esse indicador, ele vai ser o que?

[05:57] Ele vai ser uma divisão aqui entre o campo área, sobre o valor e esse campo aqui é um indicador multiplicado por 100, só para facilitar a visualização. Então, eu estou criando esse campo indicador na minha base `casas_f2` e eu vou deixar na `casas_f2` também, ainda na mesma base.

[06:26] Eu vou executar essa linha, “`Ctrl + Enter`”, executei. Aqui, eu vou deixar só para a gente visualizar `casas_f2`, “`Ctrl + Enter`” aqui. Então, se a gente passar aqui a base de dados, clicar aqui do lado, a gente vai ver esse indicador criado. Então, ele cria exatamente quando? Fazendo essa conta aqui de área, sobre valor.

[06:53] Então, aqui, por exemplo, a primeira linha, 450 de área, o primeiro imóvel, sobre nove mil de aluguel e aí, esse indicador vai ser igual a cinco. Então, agora o que que a gente vai fazer? A gente vai criar uma condição aqui na base, um outro indicador, que é um indicador de visita, que a área de planejamento pediu para a gente criar.

[07:14] Esse indicador de visita vai ser o seguinte, quando esse indicador aqui que foi criado de área sobre valor, for maior que cinco, a área de planejamento, ela vai visitar o imóvel, quando for menor que cinco, ela nem vai visitar. Então, vamos criar de novo esse novo indicador.

[07:32] Então, vou chamar aqui de `casas_finais`, agora o que que eu estou fazendo? Eu vou criar uma nova base chamada `casas_finais`, uso lá o comando `mutate`, porque agora, de novo, estou criando um novo campo, aponto para a base aonde eu estou puxando esses... aonde eu vou criar esse indicador, que é `casas_f2`.

[07:54] E agora, eu vou criar esse outro indicador chamado de visitas e aí, ele é uma condição, tem a condição quando ele for maior que... ou menor que cinco. Então, a gente vai usar um outro comando aqui que é o `ifelse`, esse comando `ifelse` é utilizado justamente para... é um condicional.

[08:17] Então, quando esse `ifelse` for... `ifelse`, abre e fecha aqui, quando o meu indicador, que foi esse... `for` maior do que cinco, aí, vírgula, que é equivalente a um `então` e `senão`, condição contrária, é o `não ok`. Então, eu criei de novo esse novo indicador, quando maior que cinco que era aquela razão de área sobre `for` maior que cinco, ele vai visitar, senão não visita.

[09:08] Vou executar essa linha que a gente acabou de criar de `casas_finais`, executei, não deu nenhum erro, depois eu vou só chamar aqui a base de `casas_finais`. Aí, a gente criou esse novo indicador de visitas. Então, quando for maior que cinco, exatamente, estritamente maior que cinco, visita, senão, não visita.

[09:37] Essa primeira linha aqui, esse indicador deu exatamente cinco, como a gente colocou a condição só de maior é `não ok`, não vai visitar e assim por diante. Essa linha aqui abaixo deu 5.3, visita. Só para facilitar o retorno para a gente devolver então essa lista para a área de planejamento, o que que a gente vai fazer agora?

[10:00] A gente poderia... vamos filtrar, a gente já sabe utilizar o `Filter` aqui. então eu vou filtrar a minha base `casas_finais`, com a condição de visitas. Então, eu devolvo essa relação para a área de planejamento, área demandante, com a relação só das casas que ela tem que visitar.

[10:27] Então, a gente tem aqui oito imóveis com a condição daquele indicador que a gente acabou de criar, que facilita a tomada de decisão e as visitas... economia de visitar só as casas de interesse, de possíveis que seriam alugadas.

[10:43] E por final aqui, só para... um outro comando que também a gente já sabe fazer, a gente pode estar fazendo o seguinte: vamos combinar o comando filter com o select. Então, estaria interessado, invés de passar a base toda aqui, a gente pode fazer um resumo da base, a gente pode selecionar só os campos de interesse, junto com esses indicadores.

[11:14] Então, vamos fazer isso, eu vou dar um select, só lembrando, quando a gente dá select, a gente dá select em campos, quando a gente dá filter, dá filter em linhas. Então, eu vou selecionar, vou dar um select aqui nos campos de interesse, mais um filter, utilizando o filter também.

[11:33] Então, vamos fazer isso. O filter é exatamente esse filtro que a gente usou aqui em cima de casas_finais, só quero filtrar as casas que eu tenho interesse para visitar e aí, na sequência, a gente coloca os campos de interesse para resumir.

[11:56] Então, eu vou colocar aqui a área, só os campos que fizeram parte do indicador aqui, vou colocar o valor, o indicador que a gente criou e o outro indicador aqui chamado de visitas. Então, eu vou criar exatamente essas casas aqui, para a gente retornar.

[12:30] Então, vou executar. Opa, deu erro, errei aqui no indicador, colocar exatamente igual. Pronto. Então, eu tenho aqui uma lista resumida só do nosso indicador, que a gente acabou de criar. Então, eu posso combinar aqui entre comandos também filter, select, que a gente já sabe fazer, para facilitar o quê?

[12:54] Facilitar a visualização e a tomada de decisão pela demandante, o que facilitaria a decisão de visitas nos imóveis desejados. Então, a gente já está fornecendo exatamente uma informação para a área demandante. Ok? É isso aí.