

Traduzindo colunas e campos



8%

Transcrição

Nesta etapa, faremos a **tradução** as colunas e campos da **língua inglesa** para **portuguesa**.

É necessário que haja uma análise para avaliar se de fato este processo é relevante ou não para nosso projeto. Como nossa exploração de dados se dará de **forma didática**, é realmente importante traduzir.

Para que tenhamos uma visualização e divisão melhor, clicaremos em "TEXT" e depois em "CELL" com a indicação para cima. Em seguida, daremos duplo clique sobre a célula para criar um novo link chamado "# Tradução". Desta forma, temos a organização das ações e podemos minimizar a "Importando base de dados" por enquanto.

Começaremos a tradução lembrando quais são as colunas que temos através da inserção de `dados.columns` na primeira célula de "Tradução". Com "Shift + Enter", o sistema mostrará o conteúdo das colunas entre colchetes.

[VOLTAR AO TOPO](#)

Na célula seguinte, criaremos uma lista chamada `renomear` igual a chave e o valor de cada uma das colunas. Copiaremos o conteúdo dos colchetes de `Index()` e colaremos dentro das chaves de `renomear`, separando cada coluna em uma linha.

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



13.9k xp

a

Em cada linha, passaremos o valor que queremos renomear com a tradução entre aspas simples para cada ocorrência.



8%

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



13.9k xp

a

```
renomear = {  
    'total_bill' : 'valor_da_conta',  
    'tip' : 'gorjeta',  
    'dessert' : 'sobremesa',  
    'day' : 'dia_da_semana',  
    'time' : 'hora_do_dia',  
    'size' : 'total_de_pessoas'  
}
```

COPIAR CÓDIGO

Para que nosso entendimento seja facilitado, não chamaremos mais a variável `dados` por este nome, e atribuiremos todo o conteúdo de `dados`. Como se trata de um `DataFrame`, passamos para a nova variável **gorjetas**, sendo igual a `dados.rename()` recebendo as colunas `columns` igual a `renomear`.

```
gorjetas = dados.rename(columns = renomear)
```

COPIAR CÓDIGO

VOLTAR AO TOPO

Desta forma, poderemos visualizar nossos dados a partir de `gorjetas`. Na célula seguinte, passamos o comando `gorjetas.head(1)` para visualizar a primeira linha de nossa tabela já com as nomenclaturas das colunas traduzidas.

Porém, os conteúdos dos campos continuam em inglês, como a `sobremesa` contendo *"No"* ou *"Yes"* e `hora_do_dia` com *"Dinner"* ou *"Lunch"*, por exemplo.

Começaremos com `sobremesa` utilizando `gorjetas.sobremesa`, o qual é um tipo categórico de "sim" ou "não". Para visualizarmos todos os tipos que aparecem para este campo, inserimos o comando `.unique()` à sentença.

```
gorjetas.sobremesa.unique()
```

[COPIAR CÓDIGO](#)

Assim, recebemos o retorno com os conteúdos `'No'`, `'Yes'` possíveis para este campo em específico. Copiaremos este mesmo conteúdo e criaremos uma lista do tipo `sim_ao` sendo igual a `'No'`, `'Yes'`. Em seguida, utilizaremos `:` para fazer a tradução da mesma forma como fizemos para cada coluna anteriormente.

```
sim_ao = {  
    'No' : 'Não',  
    'Yes' : 'Sim'  
}
```

[COPIAR CÓDIGO](#)[VOLTAR AO TOPO](#)

Agora, precisaremos mapear todos os `gorjetas.sobremesa` para "Sim" e "Não" utilizando o comando `map()` recebendo o valor `sim_ao`.

```
gorjetas.sobremesa.map(sim_ao)
```

[COPIAR CÓDIGO](#)

8%

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



13.9k xp

a

Executando, o sistema apresentará toda a coluna `sobremesa` com seus conteúdos ordenados pelo índice. Para ver se a alteração foi realizada com sucesso, escreveremos `gorjetas.head(1)` .

Feito isso, a resposta para este campo continua como "No" em inglês; isso acontece porque não atribuímos este mapa para nossa variável `gorjetas` . Logo, escreveremos que `gorjetas.sobremesa` é igual a `gorjetas.sobremesa.map(sim_nao)` .

```
gorjetas.sobremesa = gorjetas.sobremesa.map(sim_nao)
```

[COPIAR CÓDIGO](#)

Visualizando somente a primeira linha novamente, constataremos a alteração feita com sucesso.

Aplicaremos a mesma metodologia para o `dia_da_semana` . Veremos quais os dias da semana que aparecem como resposta para este campo usando `.unique()` novamente para `gorjetas.dia_da_semana` .

O sistema apresentará um `array()` contendo `'Sun'`, `'Sat'`, `'Thur'`, `'Fri'` entre colchetes. Em seguida, criaremos uma lista `dias` sendo igual aos mesmos valores entre chaves separados por linhas diferentes e atribuindo as traduções corretas com `:` .

[VOLTAR AO TOPO](#)

```
dias = {  
    'Sun' : 'Domingo',  
    'Sat' : 'Sábado',  
    'Thur' : 'Quinta',
```



8%

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



13.9k xp

a

```
'Fri' : 'Sexta'
```

```
}
```

[COPIAR CÓDIGO](#)

8%

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



13.9k xp

a

Da mesma maneira que fizemos com `sobremesa` , faremos o armazenamento destes valores dentro de `dia_da_semana` escrevendo que `gorjetas.dia_da_semana` é igual a `gorjetas.dia_da_semana.map()` recebendo `dias` .

Executaremos com "Shift + Enter" e visualizaremos a alteração com `gorjetas.head()` . Para termos certeza de que as ocorrências foram todas traduzidas, inseriremos `gorjetas.dia_da_semana.unique()` .

Constatada a mudança, alteraremos o `hora_do_dia` da mesma maneira. Para ver os conteúdos possíveis, escreveremos `gorjetas.hora_do_dia.unique()` , que retornam a resposta `['Diner', 'Lunch']` no `array()` .

Em seguida, criaremos `hora` atribuindo as traduções de `'Dinner'` para `'Jantar'` e `'Lunch'` para `'Almoço'` .

```
hora = {  
    'Dinner' : 'Jantar',  
    'Lunch' : 'Almoço'  
}
```

[VOLTAR AO TOPO](#)[COPIAR CÓDIGO](#)

Agora, mapearemos `gorjetas.hora_do_dia` sendo igual a `gorjetas.hora_do_dia.map()` recebendo `hora` . Após executar este comando, visualizaremos a tabela com `'gorjetas.head(244)'` para atestar as alterações em todas as colunas.



8%

Feito isso, realizamos as traduções dos campos `sobremesa` , `dia_da_semana` e `hora_do_dia` como uma primeira manipulação dos dados.

No próximo passo, começaremos nossa primeira análise e veremos se o valor da conta influencia diretamente no valor da gorjeta.

ATIVIDADES
4 de 8

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



13.9k xp

a

VOLTAR AO TOPO