O [método get\_dummies()](https://pandas.pydata.org/docs/reference/api/pandas.get_dummies.html) da biblioteca Pandas é utilizado para transformar variáveis categóricas em variáveis binárias. Abaixo estão os parâmetros do método:

1. data: **parâmetro obrigatório** que representa o conjunto de dados e contém as variáveis categóricas a serem transformadas em variáveis binárias.
2. prefix: é um parâmetro opcional utilizado para adicionar um prefixo às colunas binárias geradas pelo método get\_dummies(). Por exemplo, se você definir o prefixo como "cat\_", as colunas binárias geradas terão nomes como "cat\_1", "cat\_2", etc.
3. prefix\_sep: parâmetro opcional utilizado para definir o separador entre o prefixo e o nome original da coluna categórica. O valor padrão é "\_".
4. columns: parâmetro opcional utilizado para selecionar as colunas específicas do conjunto de dados que devem ser transformadas em variáveis binárias. Se não for especificado, todas as colunas categóricas serão transformadas.
5. drop\_first: é um parâmetro opcional utilizado para remover a primeira coluna binária gerada pelo método get\_dummies(). Isso é feito para evitar a multicolinearidade, que é uma situação em que duas ou mais variáveis independentes estão altamente correlacionadas entre si.
6. dtype: parâmetro opcional utilizado para definir o tipo de dado das colunas binárias geradas pelo método get\_dummies(). O valor padrão é "uint8".

Abaixo está um exemplo de como podemos aplicar esses parâmetros:

import pandas as pd

# Criando um DataFrame de exemplo

df = pd.DataFrame({'cor': ['vermelho', 'azul', 'verde', 'vermelho'],

'tamanho': ['pequeno', 'médio', 'grande', 'médio'],

'formato': ['quadrado', 'redondo', 'redondo', 'quadrado']})

# Transformando as colunas categóricas em variáveis numéricas binárias

df\_dummies = pd.get\_dummies(df, columns=['cor', 'tamanho'], prefix=['cor', 'tam'], prefix\_sep='-', drop\_first=True)

# Exibindo o DataFrame resultante

df\_dummies

**Saída:**

|  | **formato** | **cor-verde** | **cor-vermelho** | **tam-médio** | **tam-pequeno** |
| --- | --- | --- | --- | --- | --- |
| 0 | quadrado | 0 | 1 | 0 | 1 |
| 1 | redondo | 0 | 0 | 1 | 0 |
| 2 | redondo | 1 | 0 | 0 | 0 |
| 3 | quadrado | 0 | 1 | 1 | 0 |

Neste exemplo, o parâmetro data é o DataFrame df que contém as colunas categóricas a serem transformadas. O parâmetro columns é uma lista que contém os nomes das colunas categóricas a serem transformadas em variáveis binárias. O parâmetro prefix é uma lista que contém os prefixos a serem adicionados às colunas binárias geradas. Aqui, estamos adicionando o prefixo "cor\_" às colunas binárias geradas a partir da coluna "cor" e o prefixo "tam\_" às colunas binárias geradas a partir da coluna "tamanho".

Adicionamos também os parâmetros prefix\_sep='-' para definir o separador entre o prefixo e o nome original da coluna categórica como um hífen (-). Também adicionamos o parâmetro drop\_first=True para remover a primeira coluna binária gerada pelo método get\_dummies(). Isso é feito para evitar a multicolinearidade, que é uma situação em que duas ou mais variáveis independentes estão altamente correlacionadas entre si.

O resultado será um novo DataFrame chamado df\_dummies, que contém as colunas originais do DataFrame df, bem como as novas colunas binárias geradas pelo método get\_dummies() com um hífen (-) como separador de prefixo e a primeira coluna binária removida.