

Para saber mais

Assim como vimos no groupby, o funcionamento do método merge, é bem parecido entre Python e SQL.

É bem comum a necessidade de unir tabelas diferentes em um projeto de Data Science. Saber fazer o merge corretamente irá te ajudar a vencer essa etapa e criar melhores modelos de machine learning.

Vamos identificar os tipos de merge que podemos fazer:

Left Merge

Quando fazemos o merge utilizando “left”, estamos querendo manter as informações da esquerda, mas adicionar informações da tabela da direita, caso o mesmo ID seja identificado.

ID	left_side
1	LEFT
2	LEFT
3	LEFT
4	LEFT

ID	right_side
3	RIGHT
4	RIGHT
5	RIGHT
6	RIGHT

ID	left_side	ID	right_side
1	LEFT	Fill with NaN	
2	LEFT	Fill with NaN	
3	LEFT	3	RIGHT
4	LEFT	4	RIGHT
		5	RIGHT
		6	RIGHT

Right Merge

O merge utilizando “right”, se baseia no mesmo método do “left”, mas nesse caso, mantemos a informação da direita e adicionamos as da esquerda.

ID	left_side
1	LEFT
2	LEFT
3	LEFT
4	LEFT

ID	right_side
3	RIGHT
4	RIGHT
5	RIGHT
6	RIGHT

ID	left_side
1	LEFT
2	LEFT

ID	left_side	ID	right_side
3	LEFT	3	RIGHT
4	LEFT	4	RIGHT
Fill with NaN		5	RIGHT
Fill with NaN		6	RIGHT

Inner Merge

Quando utilizamos o “inner” merge, apenas mantemos as informações que estão tanto na tabela da esquerda quanto da direita.

ID	left_side
1	LEFT
2	LEFT
3	LEFT
4	LEFT

ID	right_side
3	RIGHT
4	RIGHT
5	RIGHT
6	RIGHT

ID	left_side	ID	right_side
1	LEFT		
2	LEFT		
3	LEFT	3	RIGHT
4	LEFT	4	RIGHT
		5	RIGHT
		6	RIGHT

Outer Merge

Ao fazer um “outer” merge, mantemos as informações dos dois lado, e preenchemos com NaN o que não tiver referência.

ID	left_side
1	LEFT
2	LEFT
3	LEFT
4	LEFT

ID	right_side
3	RIGHT
4	RIGHT
5	RIGHT
6	RIGHT

ID	left_side	ID	right_side
1	LEFT	Fill with NaN	
2	LEFT	Fill with NaN	
3	LEFT	3	RIGHT
4	LEFT	4	RIGHT
Fill with NaN		5	RIGHT
Fill with NaN		6	RIGHT

Essa explicação foi baseada nesse [tutorial \(https://www.kaggle.com/crawford/python-merge-tutorial\)](https://www.kaggle.com/crawford/python-merge-tutorial)(inglês) criado no Kaggle por Chris Crawford.