

Algoritmos e estrutura de dados I



Pandas - Parte 2

Combinação de dataframe

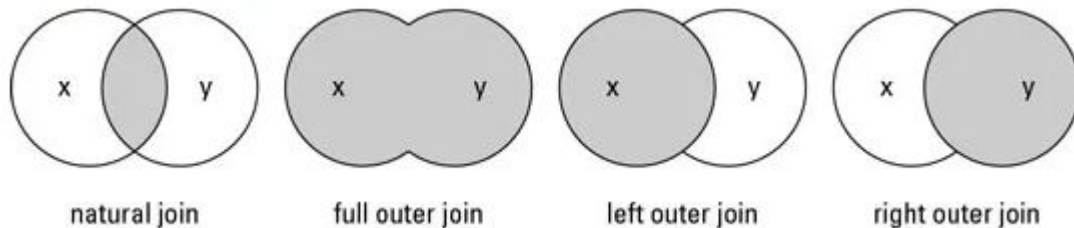
Os principais comandos para combinar dataframe são: `merge()` e `concat()`

Os métodos `merge()` e `concat()` são utilizados para unirmos data frames de maneira objetiva e com poucos dados.

Com pequenos comandos podemos unir grandes quantidades de dados.

Tipos de combinação

Assim como em bancos de dados, podemos combinar conjuntos de dados de maneiras diferentes, sendo elas:



merge()

Interseção entre os conjuntos de dados

```
>>>m = pd.merge(a, b, how = 'inner', on = 'key')
```

União entre os conjuntos de dados

```
>>>m = pd.merge(a, b, how = 'outer')
```

Conjunto da esquerda ou direita

```
>>>m = pd.merge(a, b, how = 'left', on = 'key')
```

```
>>>m = pd.merge(a, b, how = 'right', on = 'key')
```

concat()

O método concat é utilizado para concatenar linearmente o dados, ou seja, os dados serão novas linhas no dataframe.

Bastante utilizado quando o dataframe está particionado em diversos pedaços, sendo assim, ele poderá se tornar apenas um único dataframe.

```
>>>pd.concat([a_df, b_df])
```

Referências

<https://medium.com/data-hackers/pandas-combinando-data-frames-com-merge-e-concat-10e7d07ca5ec>

Livro

Título : Python para análise de dados: Tratamento de dados com Pandas, NumPy e IPython

Autor : Wes McKinney

Editora : Novatec Editora, 2019