

Merging

Merging

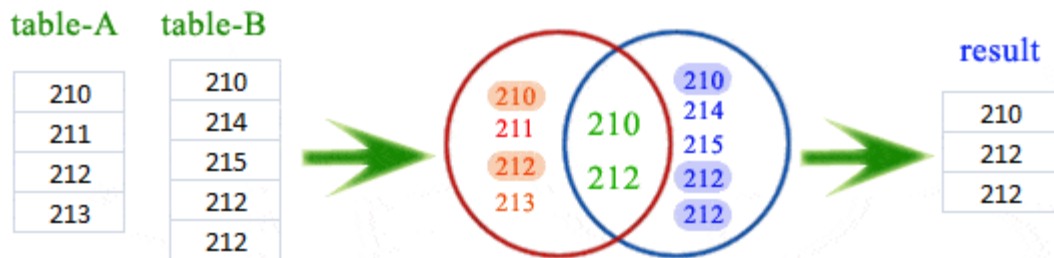
- Merging two datasets is the process of bringing two datasets together into one, and aligning the rows from each based on common attributes or columns.
- The words **merge** and **join** are used interchangeably in pandas.

Types of Joins

- There are four types of joins in python:-
 - **Inner join**
 - **Outer join**
 - **Left join**
 - **Right join**

Inner Join

- Inner join combines two dataframes based on a **common key** and returns a new dataframe that contains only rows that have matching values in both of the original dataframes.



**** matching rows being intersected**

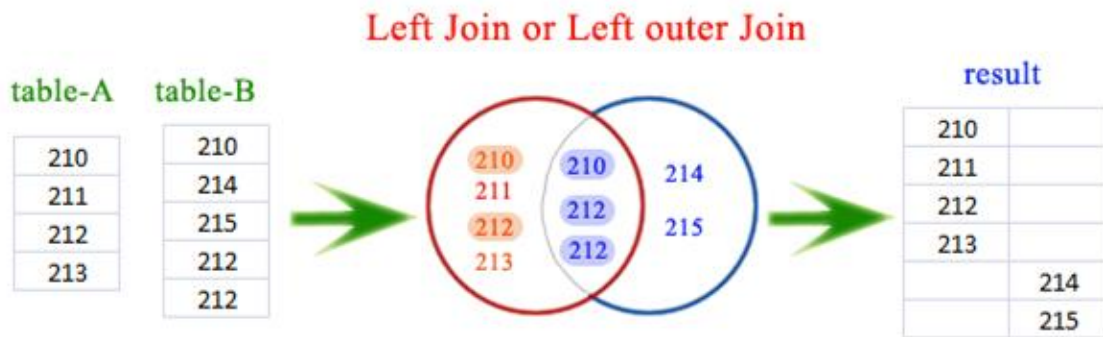
Outer Join

- Outer join returns all those records which either have a match in the left or right dataframe.
- When rows in both the dataframes do not match, the resulting dataframe will have NaN for every column of the dataframe that lacks a matching row.

TableA		TableB	
ID	Num1	ID	Num2
A	1	A	5
B	2	C	6
C	3	E	7
D	4	F	8
ID	Num1	Num 2	
A	1	5	
B	2	Null	
C	3	6	
D	4	Null	
E	Null	7	
F	Null	8	

Left Join

- Left join returns a dataframe containing all rows of left dataframe.
- All the non-matching rows of the left dataframe contain NaN for the columns in the right dataframe.



Right Join

- Right join returns a dataframe containing all rows of right dataframe.
- All the non-matching rows of the right dataframe contain NaN for the columns in the left dataframe.

TableA		TableB	
ID	Num1	ID	Num2
A	1	A	5
B	2	C	6
C	3	E	7
D	4	F	8
ID	Num1	Num 2	
A	1	5	
C	3	6	
E	Null	7	
F	Null	8	