MERGE

* 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.

In [1]:
```
import pandas as pd
```
executed in 454ms, finished 10:49:23 2021-10-26

In [35]:
```
employee_details_1 = pd.DataFrame({'employee': ['Bob', 'Jake', 'Lisa', 'Sue'],
                     'group': ['Accounting', 'Engineering', 'Engineering', 'HR']})
employee_details_2= pd.DataFrame({'employee': ['Lisa', 'Bob', 'Jake', 'Sue'],
                     'hire_date': [2004, 2008, 2012, 2014]})
```
executed in 19ms, finished 11:27:26 2021-10-26

In [21]:
```
employee_details_1
```
executed in 8ms, finished 11:21:21 2021-10-26

Out[21]:

	employee	group
0	Bob	Accounting
1	Jake	Engineering
2	Lisa	Engineering
3	Sue	HR

In [22]:
```
employee_details_2
```
executed in 7ms, finished 11:21:22 2021-10-26

Out[22]:

	employee	hire_date
0	Lisa	2004
1	Bob	2008
2	Jake	2012
3	Sue	2014

1] Merge using on

In [23]:
```
new_employee_details_1=pd.merge(employee_details_1,employee_details_2,on='employe
new_employee_details_1
```
executed in 13ms, finished 11:21:23 2021-10-26

Out[23]:

	employee	group	hire_date
0	Bob	Accounting	2008
1	Jake	Engineering	2012
2	Lisa	Engineering	2004
3	Sue	HR	2014

2] Using left on and right on

In [52]:
```
employee_details_1 = pd.DataFrame({'employee': ['Bob', 'Jake', 'Lisa', 'Sue'],
                    'group': ['Accounting', 'Engineering', 'Engineering', 'HR']})
employee_details_2= pd.DataFrame({'name': ['Lisa', 'Bob', 'Jake', 'Sue'],
                    'hire_date': [2004, 2008, 2012, 2014]})
```
executed in 8ms, finished 11:35:29 2021-10-26

In [47]:
```
new_employee_details_2=pd.merge(employee_details_1,employee_details_2,left_on='em
new_employee_details_2
```
executed in 19ms, finished 11:31:41 2021-10-26

Out[47]:

	employee	group	name	hire_date
0	Bob	Accounting	Bob	2008
1	Jake	Engineering	Jake	2012
2	Lisa	Engineering	Lisa	2004
3	Sue	HR	Sue	2014

using drop function in merge

In [48]:
```
new_employee_details_2.drop('name',axis=1)
```
executed in 15ms, finished 11:32:17 2021-10-26

Out[48]:

	employee	group	hire_date
0	Bob	Accounting	2008
1	Jake	Engineering	2012
2	Lisa	Engineering	2004
3	Sue	HR	2014

3] Using Right index and Left index

In [49]:
```
pd.merge(employee_details_1,employee_details_2,left_index=True,right_index=True)
```
executed in 15ms, finished 11:32:18 2021-10-26

Out[49]:

	employee	group	name	hire_date
0	Bob	Accounting	Lisa	2004
1	Jake	Engineering	Bob	2008
2	Lisa	Engineering	Jake	2012
3	Sue	HR	Sue	2014

4] Using how option

In [76]:
```
fav_foods = pd.DataFrame({'name': ['Peter', 'Paul', 'Mary'],
                          'food': ['fish', 'beans', 'bread']},
                         columns=['name', 'food'])
fav_drinks= pd.DataFrame({'name': ['Mary', 'Joseph'],
                          'drink': ['wine', 'beer']},
                         columns=['name', 'drink'])
```
executed in 18ms, finished 11:52:04 2021-10-26

In [77]:
```
pd.merge(fav_foods,fav_drinks,how='inner')      # Using inner option
```
executed in 23ms, finished 11:52:13 2021-10-26

Out[77]:

	name	food	drink
0	Mary	bread	wine

In [78]:
```
pd.merge(fav_foods,fav_drinks,how='outer')   # Using outer option (which provides
```
executed in 23ms, finished 11:52:34 2021-10-26

Out[78]:

	name	food	drink
0	Peter	fish	NaN
1	Paul	beans	NaN
2	Mary	bread	wine
3	Joseph	NaN	beer

In [79]:
```
pd.merge(fav_foods,fav_drinks,how='right')  # Using right option
```
executed in 10ms, finished 11:52:49 2021-10-26

Out[79]:

	name	food	drink
0	Mary	bread	wine
1	Joseph	NaN	beer

In [80]:
```
pd.merge(fav_foods,fav_drinks,how='left')   # Using left option
```
executed in 21ms, finished 11:53:08 2021-10-26

Out[80]:

	name	food	drink
0	Peter	fish	NaN
1	Paul	beans	NaN
2	Mary	bread	wine

In [ ]:
```

```