



escola  
britânica de  
artes criativas  
& tecnologia

## Profissão Cientista de Dados Tratamento de dados 2

Merge, Concat e Append

# Merge

Left join

Right join

Full outer join

Inner join

# Left join

**df1**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_3	F	24	2700
Cliente_4	M	35	6500

**df2**

ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300
Cliente_4	1000	2000

**df1.merge( df2, on = 'ID', how = 'left' )**

ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800
Cliente_3	F	24	2700	300	300
Cliente_4	M	35	6500	1000	2000

# Right join

**df1**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_3	F	24	2700
Cliente_4	M	35	6500

**df2**

ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300
Cliente_4	1000	2000

**df1.merge( df2, on = 'ID', how = 'right' )**

ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800
Cliente_3	F	24	2700	300	300
Cliente_4	M	35	6500	1000	2000

# Left join

**df1**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_3	F	24	2700
Cliente_4	M	35	6500

**df2\_sem4**

ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300

**df1.merge( df2\_sem4, on = 'ID', how = 'left' )**

ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800
Cliente_3	F	24	2700	300	300
Cliente_4	M	35	6500	NaN	NaN

# Right join

**df1**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_3	F	24	2700
Cliente_4	M	35	6500

**df2\_sem4**

ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300

**df1.merge( df2\_sem4, on = 'ID', how = 'right' )**

ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800
Cliente_3	F	24	2700	300	300

# Inner join

**df1\_sem3**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_4	M	35	6500

**df2\_sem4**

ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300

**df1\_sem3.merge( df2\_sem4, on = 'ID', how = 'inner')**

ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800



# Full outer join

**df1\_sem3**

ID	Sexo	Idade	Salario
Cliente_1	M	27	5000
Cliente_2	F	30	7000
Cliente_4	M	35	6500

**df2\_sem4**

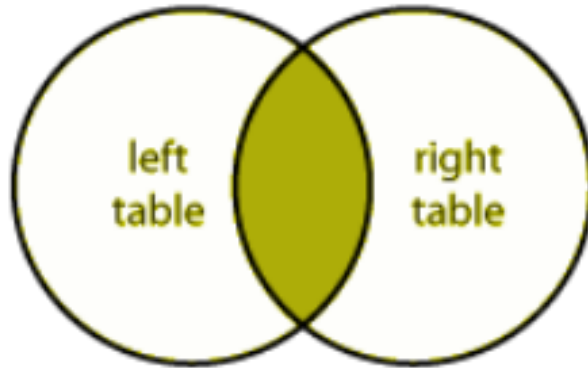
ID	Compras_M1	Compras_M2
Cliente_1	0	1000
Cliente_2	200	800
Cliente_3	300	300

**df1\_sem3.merge( df2\_sem4, on = 'ID', how = 'outer')**

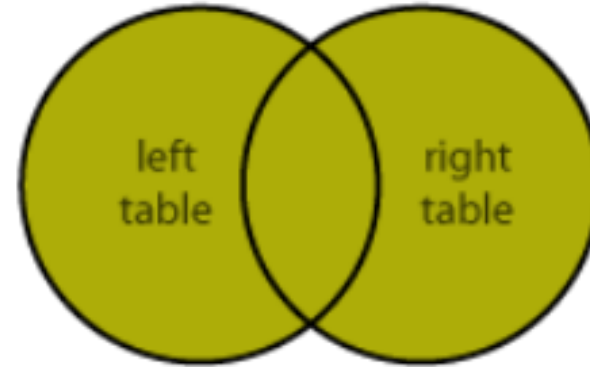
ID	Sexo	Idade	Salario	Compras_M1	Compras_M2
Cliente_1	M	27	5000	0	1000
Cliente_2	F	30	7000	200	800
Cliente_4	M	35	6500	NaN	NaN
Cliente_3	NaN	NaN	NaN	300	300

# Merge

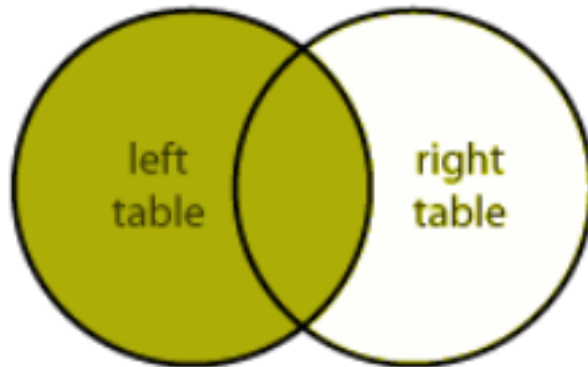
INNER JOIN



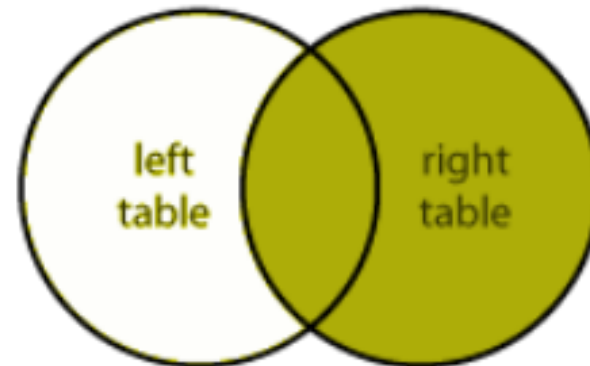
FULL JOIN



LEFT JOIN



RIGHT JOIN



# Concat

# Concat

**df1**

	ID	Sexo	Idade
0	Cliente_1	M	27
1	Cliente_2	F	30

**df2**

	ID	Sexo	Idade
0	Cliente_3	F	24
1	Cliente_4	M	35

**df3**

	ID	Sexo	Idade
0	Cliente_5	M	40
1	Cliente_6	F	50

**pd.concat ( [df1, df2, df3] , axis = 0)**

	ID	Sexo	Idade
0	Cliente_1	M	27
1	Cliente_2	F	30
0	Cliente_3	F	24
1	Cliente_4	M	35
0	Cliente_5	F	24
1	Cliente_6	M	35

# Concat

**df1**

	ID	Sexo	Idade
0	Cliente_1	M	27
1	Cliente_2	F	30

**df2**

	ID	Sexo	Idade
0	Cliente_3	F	24
1	Cliente_4	M	35

**df3**

	ID	Sexo	Idade
0	Cliente_5	M	40
1	Cliente_6	F	50

**pd.concat ( [df1, df2, df3] , axis = 1)**

	ID	Sexo	Idade	ID	Sexo	Idade	ID	Sexo	Idade
0	Cliente_1	M	27	Cliente_3	F	24	Cliente_5	M	40
1	Cliente_2	F	30	Cliente_4	M	35	Cliente_6	F	50

# Append

# Append

**df1**

	ID	Sexo	Idade
0	Cliente_1	M	27
1	Cliente_2	F	30

**df2**

	ID	Sexo	Idade
0	Cliente_3	F	24
1	Cliente_4	M	35

**df3**

	ID	Sexo	Idade
0	Cliente_5	M	40
1	Cliente_6	F	50

**df1.append( [df2, df3] )**

**pd.concat ( [df1, df2, df3] , axis = 0)**

	ID	Sexo	Idade
0	Cliente_1	M	27
1	Cliente_2	F	30
0	Cliente_3	F	24
1	Cliente_4	M	35
0	Cliente_5	F	24
1	Cliente_6	M	35