

# CHEATSHEET PANDAS – UNION E INTEGRACION DE DATOS

## APILAR (SIN CRUZAR POR NINGUNA VARIABLE)

- **Apilar variables:** `pd.concat([variables_izq, variables_drc], axis = 'columns')`

```
#Recordamos variables_izq
variables_izq
```

id	Funded Date	Funded Amount	Country	Country Code	Loan Amount	Paid Date
84	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:00:40+00:00
85	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:04:33+00:00
86	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:06:56+00:00
...	...	...	...	...	...	...
5257	2007-06-02 03:57:56+00:00	1000	Ghana	GH	1000	2008-03-13 10:01:32+00:00
5258	2007-03-24 06:01:16+00:00	1100	Ghana	GH	1100	2007-12-26 10:15:55+00:00
5259	2007-06-05 04:56:45+00:00	800	Ghana	GH	800	2008-03-13 10:01:33+00:00

5146 rows × 6 columns

```
#Recordamos variables_drc
variables_drc
```

id	Paid Amount	Activity	Sector	Delinquent	Name	Use	Status
84	500.0	Butcher Shop	Food	False	Justine	Buy bulls, open a butcher shop	paid
85	500.0	Food Production/Sales	Food	False	Geoffrey	Buying more produce each time for greater profit	paid
86	500.0	Animal Sales	Agriculture	False	Apollo	Increase number of goats bought and sold each ...	paid
...	...	...	...	...	...	...	...
5257	1000.0	Construction	Construction	False	Able	Working capital	paid
5258	1100.0	Construction	Construction	False	Manas	Working capital	paid
5259	800.0	Food Market	Food	False	Beauty	Working capital	paid

5146 rows × 7 columns

```
#Apilamos variables
pd.concat([variables_izq, variables_drc], axis = 'columns')
```

id	Funded Date	Funded Amount	Country	Country Code	Loan Amount	Paid Date	Paid Amount	Activity	Sector	Delinquent	Name	Use	Status
84	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:00:40+00:00	500.0	Butcher Shop	Food	False	Justine	Buy bulls, open a butcher shop	paid
85	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:04:33+00:00	500.0	Food Production/Sales	Food	False	Geoffrey	Buying more produce each time for greater profit	paid
86	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:06:56+00:00	500.0	Animal Sales	Agriculture	False	Apollo	Increase number of goats bought and sold each ...	paid
...	...	...	...	...	...	...	...	...	...	...	...	...	...
5257	2007-06-02 03:57:56+00:00	1000	Ghana	GH	1000	2008-03-13 10:01:32+00:00	1000.0	Construction	Construction	False	Able	Working capital	paid
5258	2007-03-24 06:01:16+00:00	1100	Ghana	GH	1100	2007-12-26 10:15:55+00:00	1100.0	Construction	Construction	False	Manas	Working capital	paid
5259	2007-06-05 04:56:45+00:00	800	Ghana	GH	800	2008-03-13 10:01:33+00:00	800.0	Food Market	Food	False	Beauty	Working capital	paid

5146 rows × 13 columns

- **Apilar registros:** `pd.concat([registros_sup, registros_inf], axis = 'index')`

```
#Recordamos registros_sup
registros_sup
```

id	Funded Date	Funded Amount	Country	Coun Co
84	2005-03-31 06:27:55+00:00	500	Uganda	
85	2005-03-31 06:27:55+00:00	500	Uganda	
86	2005-03-31 06:27:55+00:00	500	Uganda	
...	...	...	...	...
2682	2006-12-27 19:34:35+00:00	1875	Moldova	
2683	2006-12-15 22:59:04+00:00	200	Kenya	
2684	2006-12-16 15:54:26+00:00	575	Kenya	

2573 rows × 5 columns

```
#Recordamos registros_inf
registros_inf
```

id	Funded Date	Funded Amount	Country	Coun Co
2685	2006-12-15 23:46:28+00:00	375	Kenya	
2686	2007-01-10 06:14:14+00:00	2000	Bulgaria	
2687	2006-12-29 19:59:38+00:00	2000	Bulgaria	
...	...	...	...	...
5257	2007-06-02 03:57:56+00:00	1000	Ghana	
5258	2007-03-24 06:01:16+00:00	1100	Ghana	
5259	2007-06-05 04:56:45+00:00	800	Ghana	

2573 rows × 5 columns

```
#apilamos registros
pd.concat([registros_sup, registros_inf], axis = 'index')
```

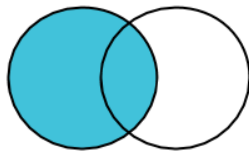
id	Funded Date	Funded Amount	Country	Coun Co	Funded Date	Funded Amount	Country	Coun Co
86	2005-03-31 06:27:55+00:00	500	Uganda	UG	500	2005-12-13 12:06:56+00:00		
...	...	...	...	...	...	...	...	...
5257	2007-06-02 03:57:56+00:00	1000	Ghana	GH	1000	2008-03-13 10:01:32+00:00		
5258	2007-03-24 06:01:16+00:00	1100	Ghana	GH	1100	2007-12-26 10:15:55+00:00		
5259	2007-06-05 04:56:45+00:00	800	Ghana	GH	800	2008-03-13 10:01:33+00:00		

5146 rows × 9 columns

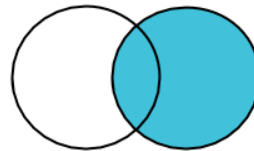
# CHEATSHEET PANDAS – UNION E INTEGRACION DE DATOS

## CRUZAR FICHEROS POR CAMPO CLAVE

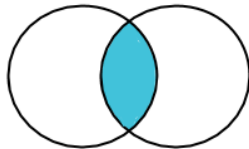
• `pd.merge (left = fichero1, right = fichero2, how = 'inner', on = 'Clave')`



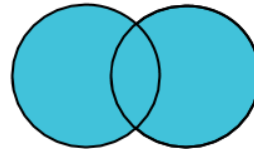
**Left Join**



**Right Join**



**Inner Join**



**Full Outer  
Join**

• **Tipos de unión (parámetro how):**

- **left:** manda la izquierda, es decir se incluirán todos los registros de la tabla de la izquierda independientemente de que estén o no en la derecha
- **right:** manda la derecha, es decir se incluirán todos los registros de la tabla de la derecha independientemente de que estén o no en la izquierda
- **inner:** solamente se incluirán los registros que estén en ambas tablas
- **outer:** se incluirán todos los registros tanto los de la izquierda como los de la derecha