Very basic pandas

https://github.com/pandas-dev/pandas/blob/master/doc/cheatsheet/Pandas_Cheat_Sheet.pdf (https://github.com/pandas-dev/pandas/blob/master/doc/cheatsheet/Pandas_Cheat_Sheet.pdf)

- load // save
- dummies
- label encoder
- Select colums
- add head
- apply function

Lee el ficher ex.csv y añadele el header

In [1]:

Guárdalo de nuevo como ex2.csv y vuélvelo a cargar

In [83]:

Out[83]:

	Name	Sex	Age	Country	Level
0	Miguel	Hombre	34.0	España	A
1	Javier	Hombre	23.0	Brasil	B
2	Ana	Mujer	40.0	Italia	A
3	Pablo	Hombre	NaN	USA	C
4	Eva	Mujer	23.0	Italia	A
5	Esteban	Hombre	-2.0	Argentina	B

Elimina las filas con valores nulos

In [84]:

Out[84]:

	Name	Sex	Age	Country	Level
0	Miguel	Hombre	34.0	España	A
1	Javier	Hombre	23.0	Brasil	B
2	Ana	Mujer	40.0	Italia	A
4	Eva	Mujer	23.0	Italia	A
5	Esteban	Hombre	-2.0	Argentina	B

Convierte la columna Sex en 0s y 1s

Usa label encoder

In [85]:

```
/Library/Python/2.7/site-packages/ipykernel/__main__.py:4: SettingWithCop
yWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy (http://pandas.pydata.or
g/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
```

Out[85]:

	Name	Sex	Age	Country	Level
0	Miguel	0	34.0	España	A
1	Javier	0	23.0	Brasil	B
2	Ana	1	40.0	Italia	A
4	Eva	1	23.0	Italia	A
5	Esteban	0	-2.0	Argentina	B

Filtra las personas con edades negativas

In [86]:

Out[86]:

	Name	Sex	Age	Country	Level
0	Miguel	0	34.0	España	A
1	Javier	0	23.0	Brasil	B
2	Ana	1	40.0	Italia	A
4	Eva	1	23.0	Italia	A

Crea una columna con la longitud de los nombres

In [87]:

```
/Library/Python/2.7/site-packages/ipykernel/__main__.py:4: SettingWithCop
yWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-doc
s/stable/indexing.html#indexing-view-versus-copy (http://pandas.pydata.or
g/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
```

Out[87]:

	Name	Sex	Age	Country	Level	NameLen
0	Miguel	0	34.0	España	A	6
1	Javier	0	23.0	Brasil	B	6
2	Ana	1	40.0	Italia	A	3
4	Eva	1	23.0	Italia	A	3

Crea una columna diciendo si son hemisferio sur o norte

In [88]:

```
/Library/Python/2.7/site-packages/ipykernel/__main__.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy (http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy)
```

Out[88]:

	Name	Sex	Age	Country	Level	NameLen	Hemisphere
0	Miguel	0	34.0	España	A	6	1
1	Javier	0	23.0	Brasil	B	6	0
2	Ana	1	40.0	Italia	A	3	1
4	Eva	1	23.0	Italia	A	3	1

Crear dummies para Sex

https://pandas.pydata.org/pandas-docs/stable/generated/pandas.get_dummies.html (https://pandas.pydata.org/pandas-docs/stable/generated/pandas.get_dummies.html)

In [89]:

In [91]:

Out[91]:

	Sex	Age	NameLen	Hemisphere	Level_A	Level_B
0	0	34.0	6	1	1	0
1	0	23.0	6	0	0	1
2	1	40.0	3	1	1	0
4	1	23.0	3	1	1	0