

Python para Análisis de Datos

Módulo 04

Pandas

Índices

Las etiquetas son una parte fundamental de las estructuras de Pandas, muchas de las funcionalidades dependen de ellas. Por lo tanto, es importante poder manipular y modificarlas según sea necesario. Las etiquetas están alojadas en una estructura llamada **Index**. En un dataframe tenemos un Index para las filas y otro para las columnas en los atributos `index` y `columns`.

Es posible modificar un index completamente simplemente reasignándolo, por ejemplo, por una lista que debe tener la misma cantidad de elementos.



Índices

df

	0	1	2
0	3	6	2
1	6	8	5
2	7	0	5
3	4	3	8

```
df.index = ["fila 1", "fila 2", "fila 3", "fila 4"]  
df.columns = ["col 1", "col 2", "col 3"]
```

df

	col 1	col 2	col 3
fila 1	3	6	2
fila 2	6	8	5
fila 3	7	0	5
fila 4	4	3	8

Índices

Los objetos de tipo `Index` no soportan la asignación de ítems, no se puede cambiar uno de sus elementos por asignación directa. Esta modificación la podemos realizar de dos maneras:

Una es modificando directamente el array de `numpy` que aloja los datos del `Index` a través del atributo `values`.

La otra forma es usando el método `rename`, que sirve para filas y columnas. Se le puede pasar un diccionario con los nombres de las etiquetas que queremos cambiar. También se puede pasar una función a aplicar a todos los elementos del `Index`. Por defecto, esta operación devuelve otro `dataframe` con los índices modificados. Si queremos modificar los índices del `dataframe` actual podemos usar el parámetro `inplace = True`.



Índices

```
df.index[0] = "FILA 1"
```

`TypeError`: Index does not support mutable operations

```
df.index.values[0] = "FILA 1"  
df
```

	col 1	col 2	col 3
FILA 1	3	6	2
fila 2	6	8	5
fila 3	7	0	5
fila 4	4	3	8

```
df.rename(index={"fila 2": "FILA 2"}, inplace=True)  
df
```

	col 1	col 2	col 3
FILA 1	3	6	2
FILA 2	6	8	5
fila 3	7	0	5
fila 4	4	3	8

```
df.rename(index = str.upper,  
          columns = lambda x: x.replace("col", "Columna"),  
          inplace = True)  
df
```

	Columna 1	Columna 2	Columna 3
FILA 1	3	6	2
FILA 2	6	8	5
FILA 3	7	0	5
FILA 4	4	3	8

reset_index

Para convertir el Index en una columna se usa el método `reset_index`. Éste transforma el Index en un columna y resetea el índice a un rango de números.

```
df
```

	Articulos vendidos	Facturación
Mes		
Enero	25	34700
Febrero	32	41500
Marzo	29	37200

Si solamente queremos resetear el Index y descartar los valores sin convertirlos en una columna, usamos el parámetro `drop = True`.

```
df.reset_index()
```

	Mes	Articulos vendidos	Facturación
0	Enero	25	34700
1	Febrero	32	41500
2	Marzo	29	37200

reindex

El método `reindex` permite cambiar el orden en que se encuentran las etiquetas pero manteniendo la alineación de los datos con la etiqueta original. Sólo se conservan las filas cuyas etiquetas se indiquen explícitamente.

Si se agregan etiquetas nuevas, los valores correspondientes se llenan con missing values a menos que explícitamente se indique el valor deseado en el parámetro `fill_value`.



reindex

```
df
```

	Articulos vendidos	Facturación
Mes		
Enero	25	34700
Febrero	32	41500
Marzo	29	37200

```
df.reindex(["Marzo", "Enero", "Septiembre"])
```

	Articulos vendidos	Facturación
Mes		
Marzo	29	37200
Enero	25	34700
Septiembre	NaN	NaN

¡Muchas gracias!

¡Sigamos trabajando!