


**Autores:** María Inmaculada Rodríguez García , María Gema Carrasco García, Javier González Enrique, Juan Jesús Ruiz Aguilar, Ignacio J. Turias Domínguez. [Universidad de Cádiz](#)

## 2.3. Funciones para obtener información

En **Python**, agrupar el número de apariciones de un valor de una columna:

```
df ['columna'].value_counts()
Número de filas del dataframe: len(df)
Número de valores distintos de una columna: df['columna'].nunique()
```

```
% En Matlab podríamos hacerlo así:
% Ejemplo de una matriz de datos con una columna
datos = [1; 2; 3; 2; 1; 3; 1; 4; 2; 3];

% Contar apariciones de cada valor en la columna
conteo_valores = tabulate(datos);

% Mostrar los resultados
disp(conteo_valores);

% Ejemplo de una matriz de datos aleatorios
datos = rand(100, 3); % Crear una matriz de 100 filas y 3 columnas con valores
aleatorios

% Obtener el número de filas (observaciones)
num_filas = size(datos, 1);

% Mostrar el número de filas
disp(num_filas);

% Obtener valores únicos en la columna
valores_unicos = unique(datos);


% Contar el número de valores únicos
num_valores_unicos = numel(valores_unicos);

% Mostrar el número de valores únicos
disp(num_valores_unicos);
```

### Ejercicio 73\_01. Funciones para obtener información en Python

```
import pandas as pd

df=pd.DataFrame({"Est": ["A", "B", "C", "D"], "Contaminante 1": [30, 50, 20, 12],
"Contaminante 2": [14, 12, 7, 4], "Tipo": ["Train", "Test", "Train", "Val"]})
print(df)
print("Filas de df: ")
print(len(df))
print("Veces que aparecen los valores de la columna tipo: ")
print(df['Tipo'].value_counts())
print('Número de tipos distintos: ')
print(df['Tipo'].value_counts())
print("Sumario general: ")
print(df.describe())
```

**Autores:** María Inmaculada Rodríguez García , María Gema Carrasco García, Javier González Enrique, Juan Jesús Ruiz Aguilar, Ignacio J. Turias Domínguez. [Universidad de Cádiz](#)

```
DataTable:
   Est  Contaminante1  Contaminante2  Tipo
0    A             30             14  Train
1    B             50             12   Test
2    C             20              7  Train
3    D             12              4   Val
```

Filas de df: 4

Veces que aparecen los valores de la columna tipo:

Train 2

Test 1

Val 1

Name: Tipo, dtype: int64

Número de tipos distintos:

3

Sumario general:

	Contaminante1	Contaminante2
count	4.000000	4.000000
mean	28.000000	18.000000
std	16.411378	9.250000
min	12.000000	4.573474
25%	18.000000	4.000000
50%	25.000000	6.250000
75%	35.000000	9.500000
Max	50.000000	14.000000

### Ejercicio 73\_02. Funciones para obtener información en MATLAB

```
% Creamos un DataTable en MATLAB
```

```
estaciones = {'A','B','C','D'};
tipos = {'Train','Test','Train','Val'};
```

```
cont1 = [30 50 20 12];
cont2 = [14 12 7 4];
```

```
nombres_columnas = {'Est','Contaminante1','Contaminante2','Tipo'};
df2 = table(estaciones,cont1,cont2,tipos,'VariableNames',nombres_columnas);
df = table(data(:, 1), data(:, 2), data(:, 3), data(:, 4), 'VariableNames',
nombres_columnas);
```

```
% Mostramos el DataTable
```


```
disp('DataTable:');
disp(df);
```

```
% Número de filas en df
```

```
num_filas = size(df, 1);
fprintf('Filas de df: %d\n', num_filas);
```

```
% Veces que aparecen los valores de la columna 'Tipo'
```

```
conteo_tipo = countcats(categorical(df.Tipo));
fprintf('Número de tipos distintos: %d\n', numel(conteo_tipo));
```

**Autores:** María Inmaculada Rodríguez García , María Gema Carrasco García, Javier González Enrique, Juan Jesús Ruiz Aguilar, Ignacio J. Turias Domínguez. [Universidad de Cádiz](#)

```
% Sumario general
disp('Sumario general:');
s = summary(df);
disp(s);
disp('Sumario de un contaminante');
disp(s.Contaminante1);

Sumario general:
      Est: [1x1 struct]
Contaminante1: [1x1 struct]
Contaminante2: [1x1 struct]
      Tipo: [1x1 struct]

Sumario de un contaminante:

struct with fields:

      Size: [4 1]
      Type: 'double'
Description: ''
      Units: ''
Continuity: []
      Min: 12
      Median: 25
      Max: 50
NumMissing: 0

DataTable:
      Est      Contaminante1      Contaminante2      Tipo
      ----      -
      {'A'}      {[30]}      {[14]}      {'Train'}
      {'B'}      {[50]}      {[12]}      {'Test' }
      {'C'}      {[20]}      {[ 7]}      {'Train'}
      {'D'}      {[12]}      {[ 4]}      {'Val'  }

Filas de df: 4
Número de tipos distintos: 3
Sumario general:
      Est: [1x1 struct]
Contaminante1: [1x1 struct]
Contaminante2: [1x1 struct]
      Tipo: [1x1 struct]
```