

# 1 - 2

Si tenemos un dataframe df con columnas age, salary, sex y queremos:

1. Mostrar aquellos individuos menores de 30 años o con salario igual o superior a 25000 euros,
2. Añadir una nueva columna numérica al dataframe llamada salary\_irpf que se corresponda con el 20% de su sueldo, entonces escribiremos:
3. Si tenemos un dataframe df con columnas age, salary, sex y queremos saber cuál es la media de edad de nuestros individuos, entonces escribiremos:
4. Si tenemos un dataframe df con columnas age, salary, sex y queremos borrar del mismo la columna sex, entonces escribiremos:
5. Si tenemos un dataframe df con columnas age, salary, sex y queremos conocer el salario máximo por sexo, entonces escribiremos:
6. Si tenemos un dataframe df con columnas age, salary, sex y queremos recuperar las 3 primeras filas y las 3 últimas de la columna salary en una única instrucción, entonces escribiremos:
7. Si tenemos un dataframe df con columnas age, salary, sex y queremos añadir borrar todas las filas que tengan valores nulos en la columna salary, entonces escribiremos:

## - - 2.

```
In [64]: import pandas as pd

# Paso 1: Cargar el archivo CSV en un DataFrame

df = pd.read_csv('demo-examen-parte2.csv')
```

Mostrar individuos menores de 30 años o con salario igual o superior a 25000 euros

```
In [72]: filtro_edad_salario = (df['age'] < 30) | (df['salary'] >= 25000)
resultado_paso_1 = df[filtro_edad_salario]
```

```
In [75]: resultado_paso_1
```

```
Out[75]:
```

|   |    |       |        |
|---|----|-------|--------|
| 0 | 58 | 36000 | male   |
| 1 | 23 | 18000 | female |
|   | 28 | 40000 | female |
|   | 46 | 50000 | male   |
|   | 18 | 15000 | male   |
|   | 38 | 38000 | female |
|   | 40 | 29000 | male   |

Añadir una nueva columna salary\_irpf con el 20% de sueldo

```
In [78]: df['salary_irpf'] = df['salary'] * 0.2  
df
```

```
Out[78]:
```

|   |    |       |        |         |
|---|----|-------|--------|---------|
| 0 | 58 | 36000 | male   | 7200.0  |
| 1 | 23 | 18000 | female | 3600.0  |
| 2 | 31 | 23000 | female | 4600.0  |
|   | 28 | 40000 | female | 8000.0  |
|   | 46 | 50000 | male   | 10000.0 |
|   | 18 | 15000 | male   | 3000.0  |
|   | 38 | 38000 | female | 7600.0  |
|   | 40 | 29000 | male   | 5800.0  |

Conocer el salario máximo por sexo

In [80]:

```
salario_maximo_por_sexo = df.groupby('sex')['salary'].max()  
salario_maximo_por_sexo
```

Out[80]:

```
sex  
female    40000  
male      50000  
Name: salary, dtype: int64
```

Calcular la media de edad de los individuos

In [48]:

```
media_edad = df['age'].mean()  
media_edad
```

Out[48]: 35.25

Borrar la columna 'sex'

In [82]:

```
df = df.drop(columns=['sex'])  
df
```

KeyboardInterrupt

Recuperar las 3 primeras y 3 últimas filas de la columna 'salary'

```
In [97]: primeras_ultimas_salario = pd.concat([df['salary'].iloc[:3], df['salary'].iloc[-3:]], ignore_index=True)
primeras_ultimas_salario
```

```
Out[97]: 0    36000
         1    18000
         2    23000
         3    15000
         4    38000
         5    29000
         Name: salary, dtype: int64
```

Borrar filas con valores nulos en la columna 'salary'

```
In [98]: df_sin_nulos = df.dropna(subset=['salary'])
df_sin_nulos
```

```
Out[98]:
```

|          |    |       |         |
|----------|----|-------|---------|
| <b>0</b> | 58 | 36000 | 7200.0  |
| <b>1</b> | 23 | 18000 | 3600.0  |
| <b>2</b> | 31 | 23000 | 4600.0  |
|          | 28 | 40000 | 8000.0  |
|          | 46 | 50000 | 10000.0 |
|          | 18 | 15000 | 3000.0  |
|          | 38 | 38000 | 7600.0  |
|          | 40 | 29000 | 5800.0  |

In [104]:

```
print("Paso 1:")
print(resultado_paso_1)

print("\nPaso 2:")
print(df)

print("\nPaso 3:")
print("La media de edad es:", media_edad)

print("\nPaso 4:")
print(df)

print("\nPaso 5:")
print(df)

print("\nPaso 6:")
print(salario_maximo_por_sexo)

print("\nPaso 7:")
print(primeras_ultimas_salario)

print("\nPaso 8:")
print(df_sin_nulos)
```

Paso 1:

|   | age | salary | sex    |
|---|-----|--------|--------|
| 0 | 58  | 36000  | male   |
| 1 | 23  | 18000  | female |
| 3 | 28  | 40000  | female |
| 4 | 46  | 50000  | male   |
| 5 | 18  | 15000  | male   |
| 6 | 38  | 38000  | female |
| 7 | 40  | 29000  | male   |

Paso 2:

|   | age | salary | salary_irpf |
|---|-----|--------|-------------|
| 0 | 58  | 36000  | 7200.0      |
| 1 | 23  | 18000  | 3600.0      |
| 2 | 31  | 23000  | 4600.0      |
| 3 | 28  | 40000  | 8000.0      |
| 4 | 46  | 50000  | 10000.0     |
| 5 | 18  | 15000  | 3000.0      |
| 6 | 38  | 38000  | 7600.0      |
| 7 | 40  | 29000  | 5800.0      |

Paso 3:

La media de edad es: 35.25

Paso 4:

|   | age | salary | salary_irpf |
|---|-----|--------|-------------|
| 0 | 58  | 36000  | 7200.0      |
| 1 | 23  | 18000  | 3600.0      |
| 2 | 31  | 23000  | 4600.0      |
| 3 | 28  | 40000  | 8000.0      |
| 4 | 46  | 50000  | 10000.0     |
| 5 | 18  | 15000  | 3000.0      |
| 6 | 38  | 38000  | 7600.0      |
| 7 | 40  | 29000  | 5800.0      |

Paso 5:

|   | age | salary | salary_irpf |
|---|-----|--------|-------------|
| 0 | 58  | 36000  | 7200.0      |
| 1 | 23  | 18000  | 3600.0      |
| 2 | 31  | 23000  | 4600.0      |
| 3 | 28  | 40000  | 8000.0      |

|   |    |       |         |
|---|----|-------|---------|
| 4 | 46 | 50000 | 10000.0 |
| 5 | 18 | 15000 | 3000.0  |
| 6 | 38 | 38000 | 7600.0  |
| 7 | 40 | 29000 | 5800.0  |

Paso 6:

sex

female 40000

male 50000

Name: salary, dtype: int64

Paso 7:

0 36000

1 18000

2 23000

3 15000

4 38000

5 29000

Name: salary, dtype: int64

Paso 8:

|   | age | salary | salary_irpf |
|---|-----|--------|-------------|
| 0 | 58  | 36000  | 7200.0      |
| 1 | 23  | 18000  | 3600.0      |
| 2 | 31  | 23000  | 4600.0      |
| 3 | 28  | 40000  | 8000.0      |
| 4 | 46  | 50000  | 10000.0     |
| 5 | 18  | 15000  | 3000.0      |
| 6 | 38  | 38000  | 7600.0      |
| 7 | 40  | 29000  | 5800.0      |