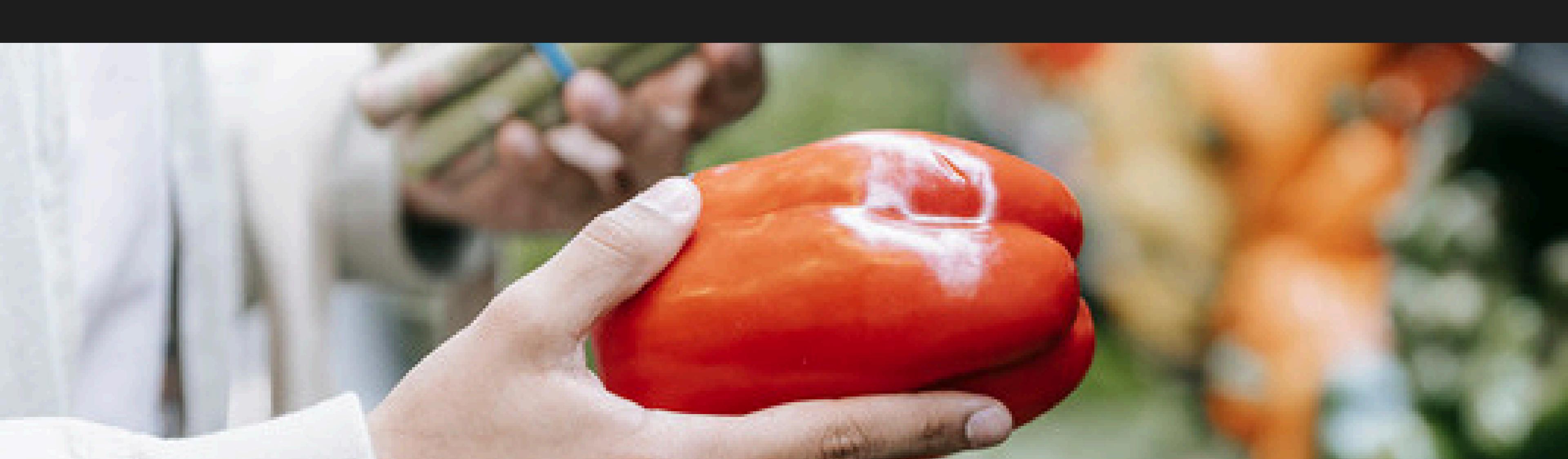




SUPERMERCADOS

PILAR ESCRIG
DANIELE BELMIRO
JAVIER RODRÍGUEZ

¿CUÁL ES EL SUPERMERCADO MÁS BARATO
PARA COMPRAR UN PRODUCTO EN
CONCRETO?



EXTRACCION DE DATOS

METODOS

Alcampo → WEB SCRAPING

Eroski → WEB SCRAPING

Dia → API's

DESAFIOS

- ENCONTRAR INFORMACION
- ACCEDER A LA INFORMACION
- DESCARGAR LA INFORMACION





LIMPIEZA Y TRANSFORMACION DE DATOS

METODOS

FIND

REPLACE

FLOAD

UNITS

DESAFIOS

-ENCONTRAR CARACTERISTICAS COMUNES

-UNIFICAR

-REEMPLAZAR

-OBTENER DETALLES ESPECIFICOS

MUESTRA DE DATOS

	Cantidad	Name	Price (€)	Uniform Price (€)	Super
37	1,5 litros	Agua mineral SOLAN DE CABRAS, botella 1,5 litros	0.89	0.59	Eroski
38	50 cl	Agua mineral EROSKI, botellín 50 cl	0.36	0.72	Eroski
39	329 ml	Agua mineral natural Dia botella 33 cl	0.24	0.73	Dia
40	33 cl	Agua mineral EROSKI, botellín tapón sport 33 cl	0.24	0.73	Eroski
41	750ml	PRODUCTO ALCAMPO Agua mineral natural de miner...	0.63	0.84	Alcampo
42	500 ml	Agua mineral natural Aquarel botella 50 cl	0.59	1.18	Dia

	Cantidad	Name	Price (€)	Uniform Price (€)	Super
0	5000 ml	Agua mineral natural Dia garrafa 5 l	0.65	0.13	Dia
1	5000 mL	Agua mineral EROSKI, garrafa 5 litros	0.74	0.15	Eroski
2	1471 ml	Agua mineral natural Dia botella 1.5 l	0.25	0.17	Dia
3	1500 mL	Agua mineral CAUTIVA, botella 1,5 litros	0.25	0.17	Eroski
4	1500 mL	Agua mineral EROSKI, botella 1,5 litros	0.39	0.26	Eroski
5	2000 mL	Agua mineral FUENTE LIVIANA, botella 2 litros	0.65	0.33	Eroski

DEMO:

```
In [ ]: 1 %load_ext autoreload  
2 %autoreload 2  
3 from Código_inicial import alcampo as alcampo  
4 from Código_inicial import eroski as eroski  
5 from Código_inicial import dia as dia  
6 import pandas as pd
```

```
In [ ]: 1 search_term = input("Introduce el término de búsqueda:")  
2 df_alcampo = alcampo(search_term)  
3 df_eroski = eroski(search_term)  
4 df_dia = dia(search_term)  
5 df_unido = pd.concat([df_alcampo, df_eroski, df_dia], axis=0)  
6 df_unido = df_unido.sort_values("Uniform Price (€)").reset_index(drop=True)  
7 df_unido
```

```
In [ ]: 1 print(f"""\n        Supermercado: {df_unido["Super"]}[0]\n        Término de búsqueda: {search_term}\n        El producto mas barato es: {df_unido["Name"]}[0]\n        Tiene {df_unido["Cantidad"]}[0]\n        Su precio uniforme es: {df_unido["Uniform Price (€)"]}[0] €\n        El precio del producto es: {df_unido["Price (€)"]}[0] €\n        """)
```

```
In [ ]: 1
```

ANALISIS DE DATOS

MÉTODOS

Concatenación: pd.concat() para unir dfs

Ordenación: sort_values() para ordenar por precio

Limpieza: replace() y strip() para consistencia

Filtrado: str.contains() para eliminar productos irrelevantes

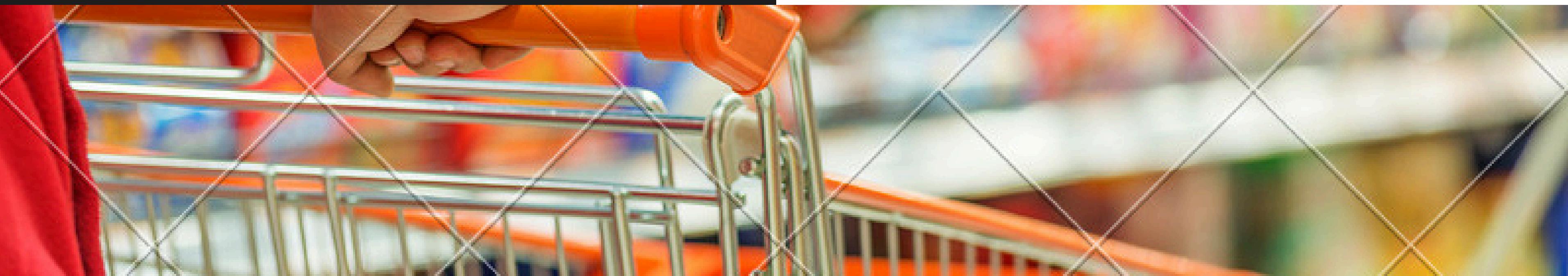
Agrupación: groupby() y value_counts() para identificar el supermercado más barato.

Cálculo de Precios Medios: mean() para precios promedio por producto.

```
1 df_huevos = df_huevos[~df_huevos['Name'].str.contains("chocolate|codorniz|pasta", case=False, na=False)]  
2
```

Pandas: manipulación de datos

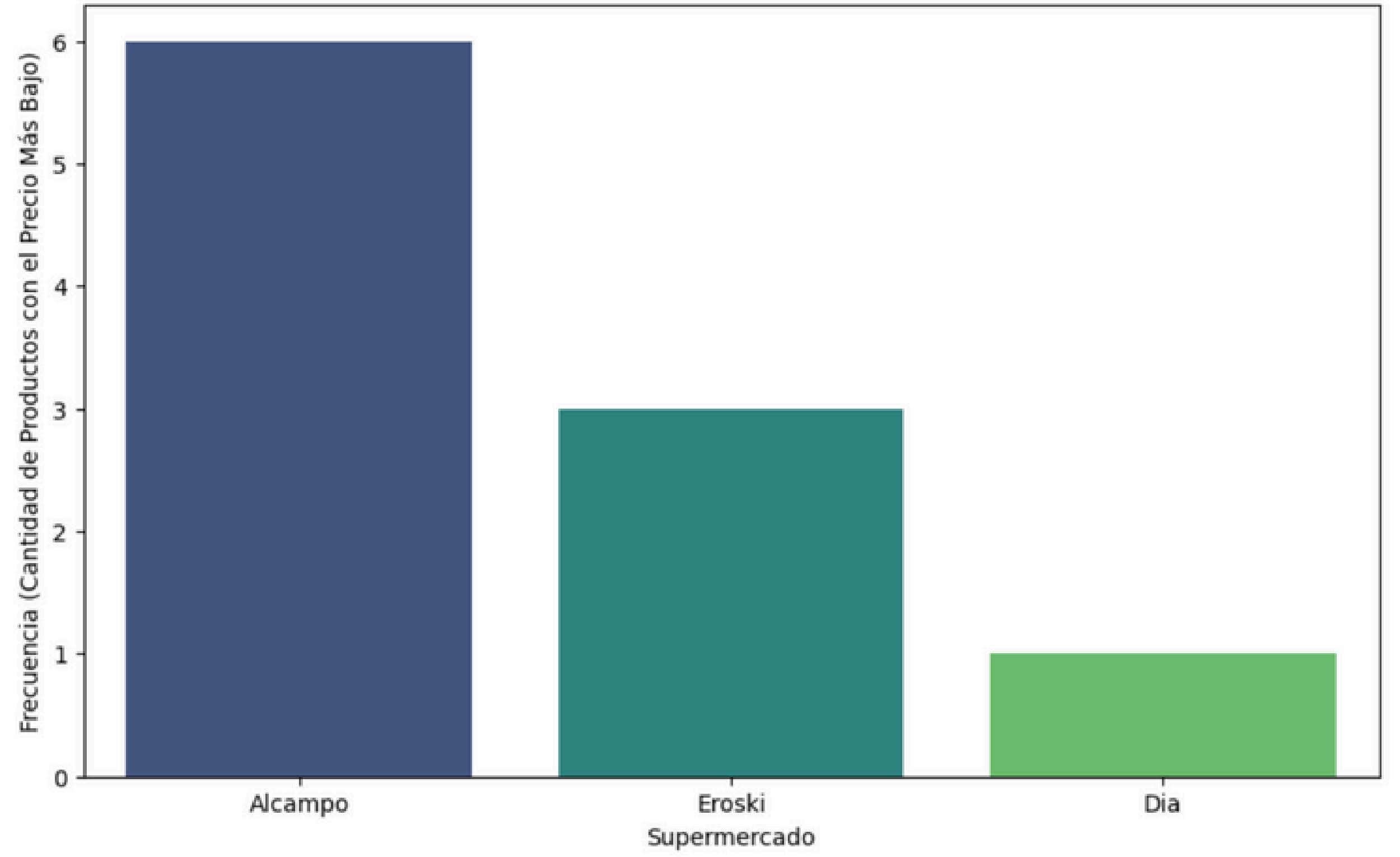
Matplotlib y Seaborn: visualización gráfica



VISUALIZACION

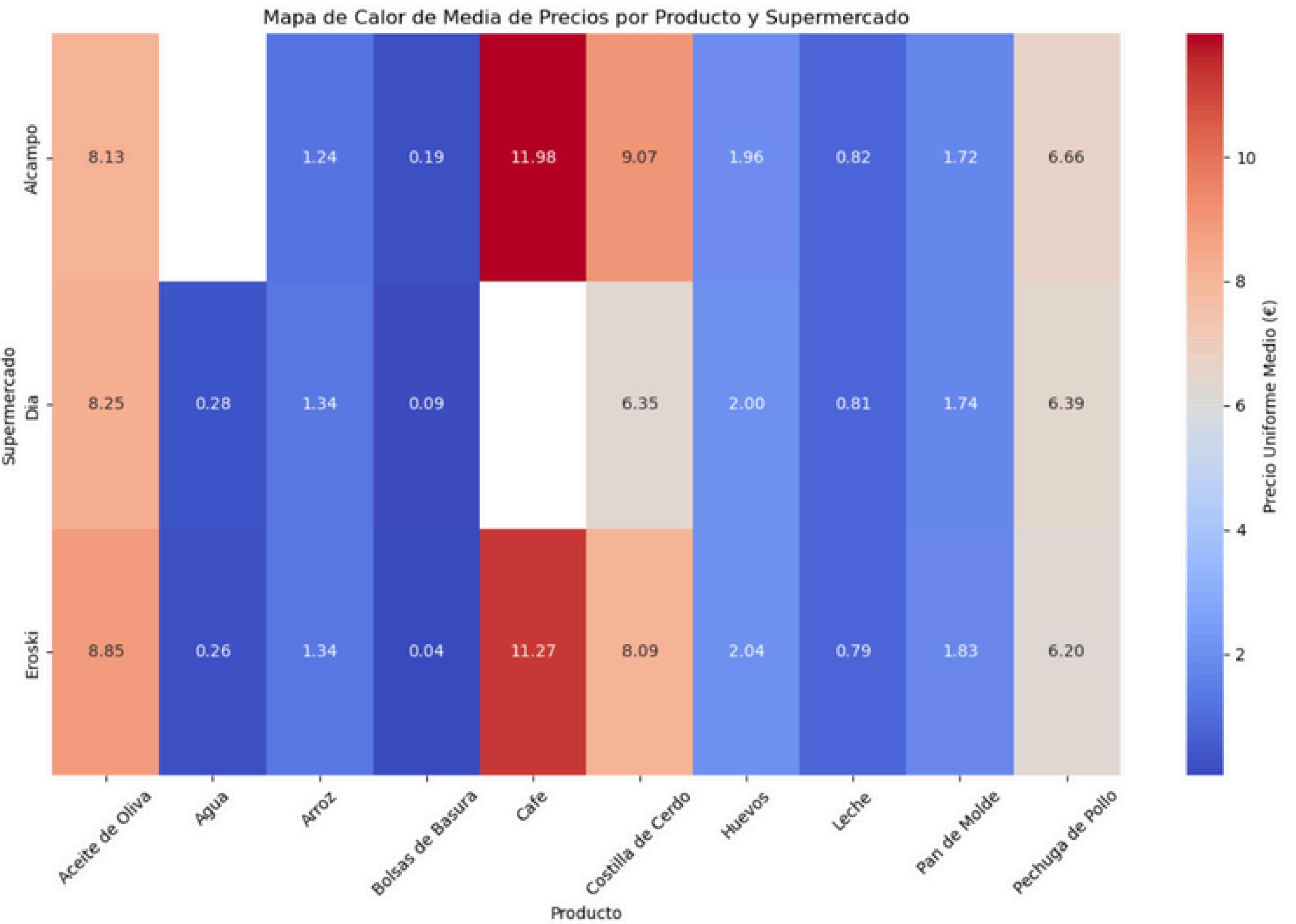
FRECUENCIA DE SUPERMERCADO MAS BARATO

Frecuencia de cada Supermercado como el Más Barato



VISUALIZACION

MAPA DE CALOR DE PRECIOS MEDIOS



PROXIMOS RETOS:

TODOS LOS SUPERMERCADOS



HIPERCOR



TODOS LOS PRODUCTOS

28	10 uds	Café en cápsulas India Nescafé Farmers Origins...	3.69	0.37	Dia
29	10 uds	Café en cápsulas Brasil Nescafé Farmers Origin...	3.69	0.37	Dia
...
56	183g	DOLCE GUSTO Café en cápsulas Latte Macchiato ...	4.88	26.67	Alcampo
57	168g	FORTALEZA INTENSSIMO Café en cápsulas I13, 2...	4.50	26.79	Alcampo
58	187g	NESCAFÉ DOLCE GUSTO Café en cápsulas con leche...	5.18	27.70	Alcampo
59	160g	DOLCE GUSTO Café en cápsulas descafeinado con ...	4.88	30.50	Alcampo
60	160g	DOLCE GUSTO Café en cápsulas con leche 16 uds.	4.88	30.50	Alcampo





Gracias