

✓ ACTIVIDAD INTEGRADORA 3: Análisis de datos con Numpy

- Pedro López
- A07777777

Leer datos de un archivo csv

```

1 import pandas as pd
2 import numpy as np
3
4 def leer_datos():
5     df = pd.read_csv("inventario.csv")
6     # Imprimo dataframe
7     # print(df)
8     # Convertir el dataframe en una matriz de numpy
9     matriz = np.array(df.values)
10    return matriz
11
12 def main():
13     m = leer_datos()
14     print(m)
15
16 main()

```

```

[[13189 'Pimienta, Florida' 228 8.75 8 1 'Condimento']
[13558 'Cuatro especias francesas' 110 6.56 2 2 'Condimento']
[15688 'Pimienta, Limón' 135 6.25 4 1 'Condimento']
[16555 'Tuscan Sunset' 95 4.55 2 2 'Condimento']
[21683 'Galena Street Rub' 125 3.95 4 1 'Adobo']
[22189 'Northwoods Fire' 135 9.85 16 1 'Condimento']
[23677 'Mejorana' 143 7.89 8 2 'Hierba']
[23688 'Curry picante en polvo' 146 9.99 8 1 'Especia']
[24896 'Butcher's Pepper' 234 5.29 4 2 'Adobo']
[25678 'Curry suave en polvo' 135 9.99 8 1 'Especia']
[25844 'Hierbas, Provence' 254 10.25 4 2 'Hierba']
[26787 'Aderezo Creole ' 165 8.75 8 2 'Condimento']
[32544 'Menta, Hierbabuena ' 156 10.29 8 2 'Hierba']
[34266 'Albahaca, Francesa' 156 5.82 8 2 'Hierba']
[34793 'Sal con cebolla' 188 3.55 2 1 'Condimento']
[34878 'Jengibre triturado' 266 7.89 8 2 'Especia']
[34982 'Charqui, pollo y pescado' 177 5.45 4 2 'Condimento']
[35677 'Charqui, puerco' 245 9.85 8 2 'Condimento']
[35690 'Charqui, Jamaica' 245 7.99 8 1 'Adobo']
[35988 'Albahaca, California' 145 11.95 8 2 'Hierba']
[36820 'Menta, Hierbabuena ' 167 10.39 8 2 'Hierba']
[37803 'Chile en polvo, picante' 248 3.39 2 1 'Condimento']
[37845 'Café' 188 17.29 8 3 'Extracto']
[38675 'Paprika húngara suave' 150 2.99 4 1 'Condimento']
[38700 'Chile en polvo, suave' 168 3.39 2 1 'Condimento']
[38744 'Bicentennial Beef' 45 4.49 4 1 'Adobo']
[39704 'Paprika, California' 133 5.79 8 1 'Condimento']
[42599 'Jengibre, cristalizado' 165 9.85 8 2 'Especia']
[43153 'Canela, China' 425 4.09 2 2 'Especia']
[43625 'Ralladura de naranja' 95 8.19 4 4 'Condimento']
[43633 'Hierbabuena' 211 5.65 4 3 'Extracto']
[43813 'Mejorana' 244 4.45 4 1 'Hierba']
[44482 'Ajo en polvo' 168 5.89 6 1 'Condimento']
[44587 'Tandoori' 75 16.85 8 2 'Especia']
[44589 'Ajo en hojuelas, California' 235 11.25 2 1 'Condimento']
[44879 'Jengibre' 160 7.95 8 1 'Especia']
[45265 'Especias para escabeche' 165 6.49 2 1 'Especia']
[45688 'Nuez moscada' 100 7.85 8 1 'Especia']
[46532 'Orégano' 265 10.19 8 1 'Hierba']
[49652 'Rojo Taco' 73 5.29 4 2 'Condimento']
[52164 'Clavos, enteros' 185 18.7 8 1 'Especia']
[53634 'Vainilla, doble intensidad' 165 16.75 8 3 'Extracto']
[54635 'Eneldo' 325 2.65 4 2 'Hierba']
[55255 'Sal marina, Pacífico' 195 2.55 8 4 'Condimento']
[56853 'Pimienta en grano, India' 312 4.59 4 1 'Especia']
[64525 'Cebolla en polvo' 152 4.85 4 1 'Condimento']
[78655 'Sal de ajo' 215 2.35 6 1 'Condimento']
[85655 'Pimiento en grano, roja' 540 3.69 2 4 'Especia']
[92258 'Vainilla' 225 15.95 4 3 'Extracto']
[93157 'Almendra' 368 7.33 4 3 'Extracto']

```

```
[93553 'Limón' 285 24.9 6 3 'Extracto']
[94236 'Comino' 126 3.55 4 2 'Especia']
[96854 'Vainilla' 423 31.95 6 3 'Extracto']
[98225 'Naranja' 325 24.19 6 3 'Extracto']
[98655 'Clavos, molidos' 211 4.55 6 1 'Especia']]
```

Envase más grande

```
1 def envase_mas_grande(m):
2     # Extraer toda la columna donde se encuentra el tamaño del envase
3     arreglo = m[:,4]
4     print(arreglo)
5     res = np.max(arreglo)
6     return res
7
8 def main():
9     m = leer_datos()
10    #print(m)
11    res = envase_mas_grande(m)
12    print("El envase más grande es de %i onzas" % res)
13
14 main()

[8 2 4 2 4 16 8 8 4 8 4 8 8 8 2 8 4 8 8 8 8 2 8 4 2 4 8 8 2 4 4 4 6 8 2 8
 2 8 8 4 8 8 4 8 4 4 6 2 4 4 6 4 6 6 6]
El envase más grande es de 16 onzas
```

Precio de menudeo más bajo

```
1 def precio_mas_bajo(m):
2     # Extraer toda la columna donde se encuentran los precios de menudeo
3     arreglo = m[:,3]
4     print(arreglo)
5     res = np.min(arreglo)
6     return res
7
8 def main():
9     m = leer_datos()
10    #print(m)
11    res = precio_mas_bajo(m)
12    print("El precio de menudeo más bajo es: ", res)
13
14 main()

[8.75 6.56 6.25 4.55 3.95 9.85 7.89 9.99 5.29 9.99 10.25 8.75 10.29 5.82
 3.55 7.89 5.45 9.85 7.99 11.95 10.39 3.39 17.29 2.99 3.39 4.49 5.79 9.85
 4.09 8.19 5.65 4.45 5.89 16.85 11.25 7.95 6.49 7.85 10.19 5.29 18.7 16.75
 2.65 2.55 4.59 4.85 2.35 3.69 15.95 7.33 24.9 3.55 31.95 24.19 4.55]
El precio de menudeo más bajo es:  2.35
```

Cantidad de productos en existencia

```
1 def suma_productos(m):
2     # Extraer toda la columna donde se encuentra la cantidad de productos en existencia
3     arreglo = m[:,2]
4     print(arreglo)
5     res = np.sum(arreglo)
6     return res
7
8 def main():
9     m = leer_datos()
10    #print(m)
11    res = suma_productos(m)
12    print("La cantidad de productos en existencia es: ", res)
13
14 main()
```

```
[228 110 135 95 125 135 143 146 234 135 254 165 156 156 188 266 177 245
245 145 167 248 188 150 168 45 133 165 425 95 211 244 168 75 235 160 165
100 265 73 185 165 325 195 312 152 215 540 225 368 285 126 423 325 211]
La cantidad de productos en existencia es: 11015
```

Promedio de los precios

```
1 def promedio_precios(m):
2     # Extraer toda la columna donde se encuentran los precios de menudeo
3     arreglo = m[:,3]
4     print(arreglo)
5     res = np.mean(arreglo)
6     return res
7
8 def main():
9     m = leer_datos()
10    #print(m)
11    res = promedio_precios(m)
12    print("El promedio de los precios de menudeo es: %.2f" % res)
13
14 main()

[8.75 6.56 6.25 4.55 3.95 9.85 7.89 9.99 5.29 9.99 10.25 8.75 10.29 5.82
3.55 7.89 5.45 9.85 7.99 11.95 10.39 3.39 17.29 2.99 3.39 4.49 5.79 9.85
4.09 8.19 5.65 4.45 5.89 16.85 11.25 7.95 6.49 7.85 10.19 5.29 18.7 16.75
2.65 2.55 4.59 4.85 2.35 3.69 15.95 7.33 24.9 3.55 31.95 24.19 4.55]
El promedio de los precios de menudeo es: 8.64
```

Cantidad de productos por categoría

```
1 def cantidad_productos(m):
2     arreglo = m[:,6]
3     # Extraer la columna de Categorías
4     #print(arreglo)
5     # Convertir el arreglo en un set
6     conjunto = set(arreglo)
7     print("Conjunto: ", conjunto)
8     #Convertir el conjunto en una lista
9     lista = list(conjunto)
10    print("Lista: ", lista)
11    lista.sort()
12    print("Lista de categorías: ", lista)
13    print("\nCategoría    #Productos")
14    for categoria in lista:
15        #print()
16        #print(categoria)
17        # Se crea una condición donde si la categoria coincide con el elemento
18        # de la columna 6
19        condicion = (m[:,6] == categoria)
20        #print(condicion)
21        # Extraer una matriz que cumpla con la condición.
22        matriz = m[condicion]
23        #print(matriz)
24        # Número de renglones de cada matriz
25        num_renglones = matriz.shape[0]
26        print(categoria + "    \t" + str(num_renglones))
27
28 def main():
29     m = leer_datos()
30     #print(m)
31     cantidad_productos(m)
32
33 main()

Conjunto: {'Extracto', 'Especia', 'Hierba', 'Adobo', 'Condimento'}
Lista: ['Extracto', 'Especia', 'Hierba', 'Adobo', 'Condimento']
Lista de categorías: ['Adobo', 'Condimento', 'Especia', 'Extracto', 'Hierba']
```

Categoría	#Productos
Adobo	4
Condimento	20
Especia	14
Extracto	8
Hierba	9

Promedio de tamaño por cada tipo de envase

```

1 def promedio_size(m):
2     envases = {1: "Frasco", 2: "Sobre", 3: "Botella", 4: "Lata"}
3     print(envases)
4     print("\nTipo de envase \t Promedio de tamaño")
5     for key in envases:
6         # print()
7         # print(key, envases[key])
8         # Condición donde coincida key con la columna de Envases (5)
9         condicion = (m[:,5] == key)
10        # Crear una matriz que cumpla con la condición
11        matriz = m[condicion]
12        #bprint(matriz)
13        arreglot = matriz[:,4]
14        # print(arreglot)
15        promedio = np.mean(arreglot)
16        print("%i. %s \t %.2f" % (key, envases[key], promedio))
17
18 def main():
19     m = leer_datos()
20     promedio_size(m)
21
22 main()

```

```
{1: 'Frasco', 2: 'Sobre', 3: 'Botella', 4: 'Lata'}
```

Tipo de envase	Promedio de tamaño
1. Frasco	5.76
2. Sobre	5.79
3. Botella	5.75
4. Lata	4.67

Cálculos para cada tipo de envase

```

1 def calculos(m, envase):
2     lista_categorias = ["Adobo", "Condimento", "Especia", "Extracto", "Hierba"]
3     matrizFin = []
4     for categoria in lista_categorias:
5         lista = []
6         #print()
7         #print(categoria)
8         # Condición donde coincida la categoria y el tipo de envase
9         condicion = (m[:,6] == categoria) & (m[:,5] == envase)
10        matriz = m[condicion]
11        #print(matriz)
12
13        #print("PRECIO MINIMO")
14        # Extraer columna 3 de los precios al menudeo
15        arreglo = matriz[:,3]
16        #print(arreglo)
17        if arreglo.size > 0:
18            preciomin = np.min(arreglo)
19        else:
20            preciomin = 0
21        #print("Precio mínimo: ", preciomin)
22        lista.append(preciomin)
23
24        #print("SUMA DE PRODUCTOS EN EXISTENCIA")
25        # Extraer columna 2 de cantidad de productos en existencia
26        arreglo = matriz[:,2]
27        #print(arreglo)
28        if arreglo.size > 0:
29            suma = np.sum(arreglo)
30        else:
31            suma = 0
32        #print("Suma de productos en existencia: ", suma)
33        lista.append(suma)
34
35        #print("TAMAÑO PROMEDIO DEL ENVASE")
36        # Extraer columna 4 del tamaño del envase
37        arreglo = matriz[:,4]
38        #print(arreglo)
39        if arreglo.size > 0:
40            promedio = np.mean(arreglo)
41        else:
42            promedio = 0
43        #print("Tamaño promedio del envase: ", promedio)
44        lista.append(promedio)
45        #print(lista)
46        matrizFin.append(lista)
47    return matrizFin
48
49 def imprime_matriz(m):
50     categorias = ["Adobo", "Condimento", "Especia", "Extracto", "Hierba"]
51     cont = 0
52     print("\t\t\tPrecio mínimo Suma producto Tamaño promedio envase")
53     for r in range(m.shape[0]): # renglones
54         print(categorias[cont], end = " ")
55         cont = cont + 1
56         for c in range(m.shape[1]): # columnas
57             print("\t %.2f" % m[r, c], end= " ")
58         print()
59
60 def main():
61     m = leer_datos()
62
63     print("\n1. FRASCO")
64     matriz = calculos(m, 1)
65     #print(matriz)
66     matriznp = np.array(matriz)

```

```

67 #print(matriznp)
68 imprime_matriz(matriznp)
69
70 print("\n2. SOBRE")
71 matriz = calculos(m, 2)
72 matriznp = np.array(matriz)
73 imprime_matriz(matriznp)
74
75 print("\n3. BOTELLA")
76 matriz = calculos(m, 3)
77 matriznp = np.array(matriz)
78 imprime_matriz(matriznp)
79
80 print("\n4. LATA")
81 matriz = calculos(m, 4)
82 matriznp = np.array(matriz)
83 imprime_matriz(matriznp)
84
85
86 main()

```

1. FRASCO

	Precio mínimo	Suma producto	Tamaño promedio envase
Adobo	3.95	415.00	5.33
Condimento	2.35	2155.00	5.33
Especia	4.55	1414.00	6.50
Extracto	0.00	0.00	0.00
Hierba	4.45	509.00	6.00

2. SOBRE

	Precio mínimo	Suma producto	Tamaño promedio envase
Adobo	5.29	234.00	4.00
Condimento	4.55	865.00	4.67
Especia	3.55	1057.00	6.00
Extracto	0.00	0.00	0.00
Hierba	2.65	1346.00	6.86

3. BOTELLA

	Precio mínimo	Suma producto	Tamaño promedio envase
Adobo	0.00	0.00	0.00
Condimento	0.00	0.00	0.00
Especia	0.00	0.00	0.00