

ACTIVIDAD INTEGRADORA 3. Análisis de datos con Numpy

- Juan Pérez Sánchez
- A00738382

Leer datos de un archivo .csv

```
In [ ]: import pandas as pd
import numpy as np

def leer_datos():
    df = pd.read_csv("inventario.csv")
    # Imprimir dataframe
    # print(df)
    # Convertir el dataframe en una matriz de numpy
    matriz = np.array(df.values)
    return matriz

def main():
    m = leer_datos()
    print(m)

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']]

```

Suma de la cantidad de productos en existencia

```
In [ ]: def suma_productos(m):  
    # Extraer la columna de cantidad en existencia  
    arreglo = m[:,2]  
    #print(arreglo)  
    res = np.sum(arreglo)  
    return res  
  
def main():  
    m = leer_datos()  
    res = suma_productos(m)  
    print("La cantidad de productos en existencia es: ", res)  
  
main()
```

La cantidad de productos en existencia es: 11015

Promedio de los precios

```
In [ ]: def promedio_precios(m):  
    # Extraer la columna de los precios  
    arreglo = m[:,3]  
    #print(arreglo)  
    res = np.mean(arreglo)  
    return res  
  
def main():  
    m = leer_datos()  
    res = promedio_precios(m)  
    print("El promedio de los precios es: %.2f" % res)  
  
main()
```

El promedio de los precios es: 8.64

Envase más pequeño

```
In [ ]: def envase_mas_chico(m):  
    # Extraer la columna de los tamaños del envase  
    arreglo = m[:,4]  
    #print(arreglo)  
    res = np.min(arreglo)  
    return res  
  
def main():  
    m = leer_datos()  
    res = envase_mas_chico(m)  
    print("El tamaño del envase más chico es de %i onzas" % res)  
  
main()
```

El tamaño del envase más chico es de 2 onzas

Precio más alto

```
In [ ]: def precio_mas_alto(m):  
    # Extraer la columna de los precios  
    arreglo = m[:,3]  
    #print(arreglo)  
    res = np.max(arreglo)
```

```

return res

def main():
    m = leer_datos()
    res = precio_mas_alto(m)
    print("El precio de menudeo más alto es: %.2f" % res)

main()

```

El precio de menudeo más alto es: 31.95

Cantidad de productos de cada categoría

```

In [ ]: def cantidad_productos(m):
        arreglo = m[:,6]
        # print(arreglo)
        # Crear un conjunto con las categorías sin repetir
        conjunto = set(arreglo)
        #print(conjunto)
        lista = list(conjunto)
        #print(lista)
        lista.sort()
        #print(lista)
        print("\nCategoría \t#Productos")
        for categoria in lista:
            # print()
            # print(categoria)
            # Crear una condición donde si la categoría coincide con el elemento
            # de la categoría 6
            condicion = (m[:,6] == categoria)
            #print(condicion)
            matriz = m[condicion]
            #print(matriz)
            num_renglones = matriz.shape[0]
            print(categoria + " \t" + str(num_renglones))

def main():
    m = leer_datos()
    cantidad_productos(m)

main()

```

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

Promedio del tamaño de cada tipo de envase

```

In [ ]: def promedio_tamaño(m):
        # Tipos de envase
        diccionario = {1: "Frasco", 2: "Sobre", 3: "Botella", 4: "Lata"}
        print("\nTipo de envase \t Promedio del tamaño")
        for key in diccionario:
            # print()
            # print(key, " : ", diccionario[key])
            # Condicion donde coincida key (Tipo de envase) con la columna 5 del tipo envase
            condicion = (m[:,5] == key)

```

```
#print(condicion)
# Crear una matriz que cumpla con la condición
matriz = m[condicion]
#print(matriz)
arreglo = matriz[:,4]
#print(arreglo)
promedio = np.mean(arreglo)
print("%i. %s \t %.2f" % (key, diccionario[key], promedio))

def main():
    m = leer_datos()
    promedio_tamaño(m)

main()
```

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