****

Міністерство освіти і науки України

Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Факультет інформатики та обчислювальної техніки

Кафедра інформаційних систем та технологій

Лабораторна робота №1

з дисципліни

Аналіз даних з використанням мови Python

|  |  |  |
| --- | --- | --- |
| Виконав: |  | Перевірила: |
|  |  |  |
| студент групи ІА-24: |  | ст. викладач |
| Криворучек В.С. |  | Тимофєєва Ю.С. |
|  |  |  |

Київ 2025

**Тема:** Базове знайомство з бібліотеками NumPy та Pandas

**Мета роботи:** Ознайомитись з основними можливостями роботи з масивами бібліотеки NumPy, структурами даних Pandas, визначити основні статистичні характеристики.

**Хід роботи**

Завдання 1. Створюємо програму, яка:  
а) генерує випадкові і невипадкові масиви різними способами, зазначеними в теоретичних відомостях.  
б) демонструє звернення до елементів масиву за допомогою індексів, в тому числі від’ємних; виділення підмасивів як одновимірних, так і багатовимірних масивів.  
в) демонструє основні арифметичні операції над масивами, а також роботу методів reduce, accumulate, outer.

Код програми:

import numpy as np  
import pandas as pd  
import os  
  
# Завдання 1  
print("\n--- Завдання 1 ---")  
  
# а) Створення різних типів масивів

print("\nа) Генерація масивів:")  
  
arr1 = np.arange(1, 16, 2)  
print("Масив із кроком 2 у діапазоні [1, 16):\n", arr1)  
  
arr2 = np.ones(9, dtype=int)  
print("Масив із одиниць:\n", arr2)  
  
arr3 = np.zeros((2, 4), dtype=float)  
print("Двовимірний масив нулів:\n", arr3)  
  
arr4 = np.linspace(0, 1, 5)  
print("Рівномірний розподіл значень від 0 до 1:\n", arr4)  
  
arr5 = np.random.rand(4, 3)  
print("Масив випадкових чисел (4x3):\n", arr5)  
  
arr6 = np.random.randint(0, 100, size=(4, 4))  
print("Масив випадкових цілих чисел (4x4):\n", arr6)  
  
arr7 = np.empty(5)  
print("Порожній масив із 5 елементів:\n", arr7)  
  
# б) Доступ до елементів масиву

print("\nб) Індексація та зрізи:")  
print("Третій рядок у масиві arr6:\n", arr6[2])  
print("Третій елемент з кінця:\n", arr6[-3])  
print("Елемент першого рядка та останнього стовпця:\n", arr6[0, -1])  
  
print("\nВиділення підмасивів:")  
print("Перші два рядки:\n", arr6[:2])  
print("Другий та третій стовпці:\n", arr6[:, 1:3])  
  
# в) Арифметичні операції

print("\nв) Операції з масивами:")  
arr8 = np.array([1, 2, 3])  
arr9 = np.array([4, 5, 6])  
  
print("Сума масивів:\n", arr8 + arr9)  
print("Добуток масивів:\n", arr8 \* arr9)  
print("Функція reduce (сума елементів arr8):\n", np.add.reduce(arr8))  
print("Функція accumulate (накопичена сума arr8):\n", np.add.accumulate(arr8))  
print("Функція outer (перемноження кожного з кожним):\n", np.multiply.outer(arr8, arr9))

Результат виконання:

--- Завдання 1 ---

а) Генерація масивів:

Масив із кроком 2 у діапазоні [1, 16):

[ 1 3 5 7 9 11 13 15]

Масив із одиниць:

[1 1 1 1 1 1 1 1 1]

Двовимірний масив нулів:

[[0. 0. 0. 0.]

[0. 0. 0. 0.]]

Рівномірний розподіл значень від 0 до 1:

[0. 0.25 0.5 0.75 1. ]

Масив випадкових чисел (4x3):

[[0.33899579 0.72964243 0.42635429]

[0.78185251 0.04645263 0.94407529]

[0.43022123 0.29121402 0.51835588]

[0.84028072 0.56597859 0.3928728 ]]

Масив випадкових цілих чисел (4x4):

[[51 53 66 7]

[42 26 49 49]

[60 26 61 58]

[69 64 98 3]]

Порожній масив із 5 елементів:

[0. 0.25 0.5 0.75 1. ]

б) Індексація та зрізи:

Третій рядок у масиві arr6:

[60 26 61 58]

Третій елемент з кінця:

[42 26 49 49]

Елемент першого рядка та останнього стовпця:

7

Виділення підмасивів:

Перші два рядки:

[[51 53 66 7]

[42 26 49 49]]

Другий та третій стовпці:

[[53 66]

[26 49]

[26 61]

[64 98]]

в) Операції з масивами:

Сума масивів:

[5 7 9]

Добуток масивів:

[ 4 10 18]

Функція reduce (сума елементів arr8):

6

Функція accumulate (накопичена сума arr8):

[1 3 6]

Функція outer (перемноження кожного з кожним):

[[ 4 5 6]

[ 8 10 12]

[12 15 18]]

Завдання 2. Варіант 10.  
а) Вивести основні статистичні характеристики кількісних ознак;  
б) Порахувати кількість та вивести інформацію щодо товарів з низьким пріоритетом та вагою до 2кг, що будуть відправлені кораблем;  
в) Додати новий стовпець, який містить перерахунок ціни за кілограм.

Код програми:

# Завдання 2  
print("\nЗавдання 2:")  
  
file\_path = 'CommerceShipping.csv'  
  
if os.path.exists(file\_path):  
 # Завантаження даних  
 data = pd.read\_csv(file\_path)  
  
 print("\nВиведення перших рядків таблиці:")  
 print(data.head().to\_string(index=False))  
  
 # Вивід статистичних характеристик  
 print("\nОсновні статистичні характеристики кількісних ознак:")  
 print(data.describe().to\_string())  
  
 # Фільтрація товарів  
 low\_prior\_under\_2kg = data[(data['Product\_importance'] == 'low') &  
 (data['Mode\_of\_Shipment'] == 'Ship') &  
 (data['Weight\_in\_gms'] < 2000)]  
  
 # Виведення загальної кількості знайдених товарів  
 print(  
 f"\nКількість товарів з низьким пріоритетом, вагою до 2кг, що будуть відправлені кораблем: {len(low\_prior\_under\_2kg)}")  
  
 # Виведення лише перших 5 рядків таблиці  
 if not low\_prior\_under\_2kg.empty:  
 print("\nПерелік таких товарів:")  
 print(low\_prior\_under\_2kg.to\_string(index=False))  
 else:  
 print("\nТоварів, що відповідають цим критеріям, немає.")  
  
 # Додаємо новий стовпець з ціною за кг  
 data['Price\_per\_kg'] = data['Cost\_of\_the\_Product'] / (data['Weight\_in\_gms'] / 1000)  
  
 print("\nТаблиця з новим стовпцем:")  
 # Перевіряємо результат (виведемо перші 5 рядків)  
 print(data.head().to\_string(index=False))  
  
else:  
 print("Файл не знайдено!")

Результат:

Зображення, що містить текст, знімок екрана, меню

Вміст, створений ШІ, може бути неправильним. Зображення, що містить текст, знімок екрана

Вміст, створений ШІ, може бути неправильним.

Зображення, що містить текст, знімок екрана, Шрифт

Вміст, створений ШІ, може бути неправильним. Зображення, що містить текст, Шрифт, знімок екрана

Вміст, створений ШІ, може бути неправильним.

**Висновок:** У ході виконання даної лабораторної роботи я зміг ознайомитись з основними можливостями роботи з масивами бібліотеки NumPy, структурами даних Pandas, визначити основні статистичні характеристики.