|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**  Дисципліна  **«Ймовірнісні основи програмної інженерії»**  **Лабораторна робота № 3**  **на тему:**  **«**ДВОВИМІРНА СТАТИСТИКА**»** | | | |
| **Виконав:** | Литвинчук Владислав Валерійович | **Перевірила**: | Вечерковська  Анастасія  Сергіївна |
| Група | ІПЗ-21(1) | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

# **Мета:**

Навчитись використовувати на практиці набуті знання про міри в двовимірній

статистиці.

# **Постановка задачі:**

1. Намалюйте діаграму розсіювання для даних. Укажіть, чи існує тренд у даних. Якщо так, то вкажіть, чи є це негативним трендом, чи позитивним.
2. Знайдіть центр ваги і коваріацію.
3. Знайти рівняння лініїї регресії y від x.
4. Розрахуйте коефіцієнт кореляції між даними.
5. Зробити висновок про залежності.

# **Математична модель:**

Для знаходження правильних відповідей для даної лабораторної роботи, необхідно використовувати формули статистики, наведені нижче:

***Центр ваги:***

де mean – середнє значення

***Коваріація:***

де n – кількість елементів у масиві значень*,*  – значення i-го елементу, mean – середнє значення масиву елементів.

***Рівняння лінії регресії y від x:***

***Коефіцієнт кореляції:***

де  **–** добуток середнього квадратичного відхилення x та y

# **Псевдокод:**

1. Знаходження центру ваги

def getCenterOfWeight(data\_matrix, data\_length):  
 знайти середнє значення суми покупок

знайти середнє значення часу, витраченого на покупку  
 центр ваги = точка G (середнє значення суми покупок, середнє значення часу)  
 return center\_of\_weight

1. Знаходження коваріації

def getCovariance(data\_matrix, data\_length):  
 знайти середнє значення суми покупок  
 знайти середнє значення часу  
  
 sum\_of\_products\_of\_deviations = 0  
 for index\_of\_data in range(data\_length):  
 девіації суми покупки = поточна сума покупки – середнє значення  
 девіації часу = поточний час – середнє значення  
 sum\_of\_products\_of\_deviations += purchase\_deviation \* time\_deviation  
  
 covariance = round(sum\_of\_products\_of\_deviations / (data\_length), 2)  
 return covariance

1. Знаходження рівняння лінії регресії y від x

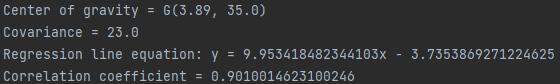
def getLinearRegression(data\_matrix, data\_length):  
 знайти коваріацію матриці значеня  
 знайти середнє квадратичне відхилення масиву суми покупок  
 b1 = covariance / variance  
 b0 = середнє значення часу – b1 \* cереднє значення суми покупок  
  
 regression\_line\_equation = f'y = {b1}x {"+" if b0 > 0 else "-"} {abs(b0)}'  
 return regression\_line\_equation

1. Знаходження коефіцієнту кореляції

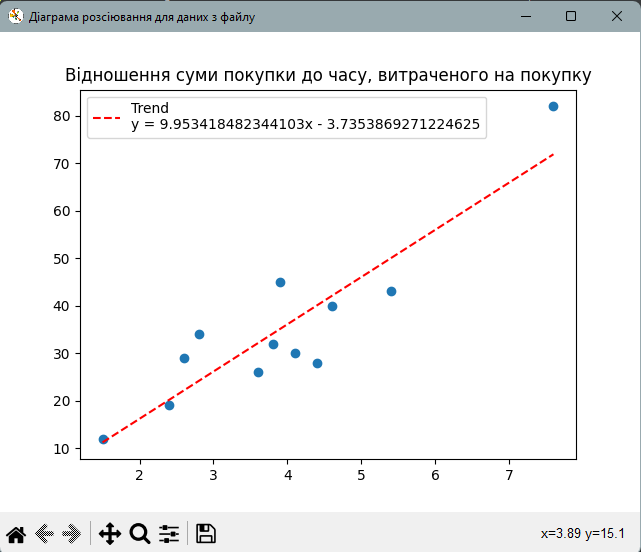
def getCorrelationCoefficient(data\_matrix, data\_length):  
 знайти коваріацію матриці значеня  
 знайти середнє квадратичне відхилення масиву суми покупок  
 знайти середнє квадратичне відхилення масиву часу  
  
 correlation\_coefficient = коваріація / середнє квадратичне X \* середнє квадратичне Y  
 return correlation\_coefficient

# **Випробування алгоритму:**

Файл output.txt



Викликана програма для відображення гістограми



# **Виcновок:**

Під час виконання даної лабораторної роботи було проведено аналіз алгоритмів і формул, необхідних для знаходження рішень. Також було розроблено алгоритми для роботи з мірами в двовимірній статистиці, а саме: побудова діаграми розсіювання даних, знаходження центру ваги і коваріативності, рівняння лінії регресії y від x, а також розрахунок коефіцієнту кореляції між даними.