

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
«Двовимірна статистика»

Виконала:	Манойлова Катерина Борисівна	Перевірила:	Вечерковська Анастасія Сергіївна
Група	ІПЗ-21	Дата перевірки	
Форма навчання	денна	Оцінка	
Спеціальність	121		

Тема: двовимірна статистика.

Мета роботи: навчитись використовувати на практиці набуті знання про міри в двовимірній статистиці.

Завдання

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 лінії регресії у від х.
4. Розрахуйте коефіцієнт кореляції між даними.
5. Зробити висновок про залежності.

Математична модель:

Центр ваги визначається як точка з координатами $G(\bar{x}; \bar{y})$

Коваріація визначається за формулою $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$

Рівняння лінії регресії: $y = b_0 + b_1 x$

$$b_1 = \frac{cov(x;y)}{var(x)}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Тренд визначається значенням коефіцієнта b_1

Коефіцієнт кореляції: $s_x = \sqrt{Var(x)}$

$$s_y = \sqrt{Var(y)}$$

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

Псевдокод алгоритмів:

Знаходження дисперсії:

```
variance():  
    sum = 0  
    for el in arr:  
        sum += (el-mid)**2  
    disp = sum/к-ть елементів
```

Знаходження коваріації:

```
covariance():  
    sum = 0  
    for x, y in arrx, arry:  
        sum += (x-середнє значення першої вибірки)*(y-середнє значення другої вибірки)  
    cov = sum/к-ть елементів
```

Знаходження коефіцієнтів лінії регресії:

linreg():

$b1 = \text{коваріація/дисперсія}(\text{arrx})$

$b0 = \text{середнє}(\text{arry}) - (b1 * \text{середнє}(\text{arrx}))$

Знаходження коефіцієнта кореляції:

corCoef():

$sx = \text{корін з дисперсії по}(\text{arrx})$

$sy = \text{корін з дисперсії по}(\text{arry})$

$\text{sum} = 0$

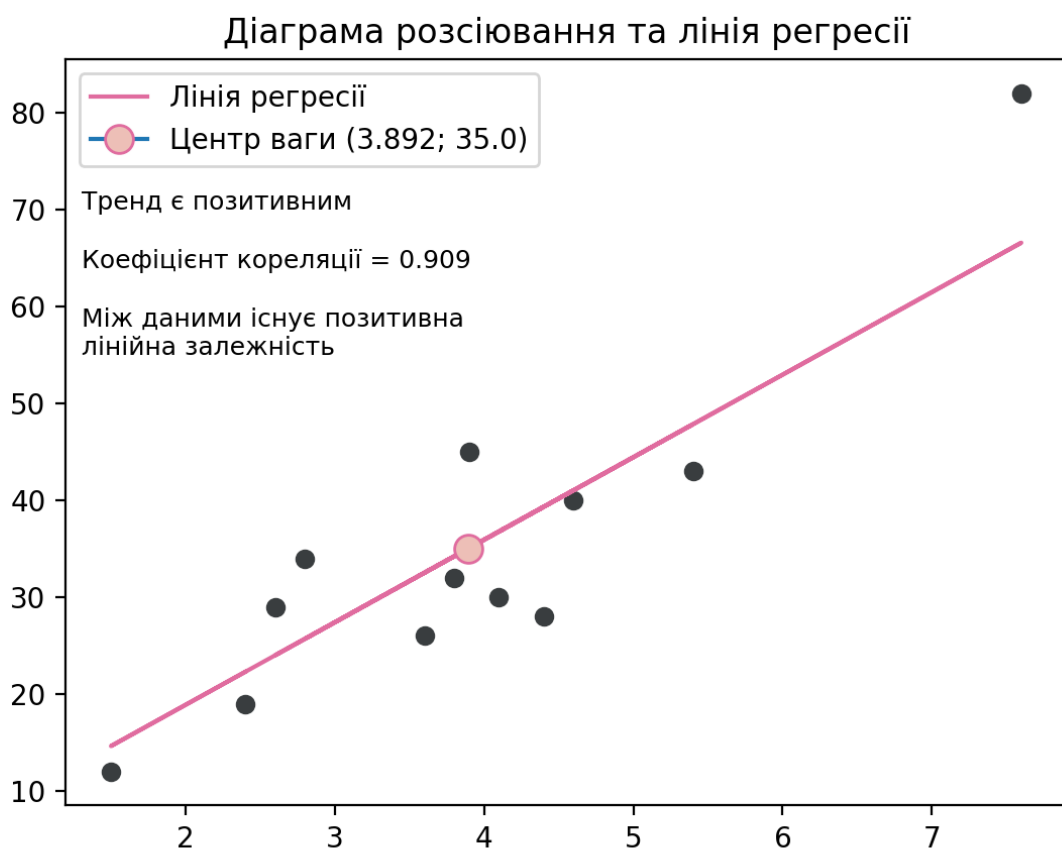
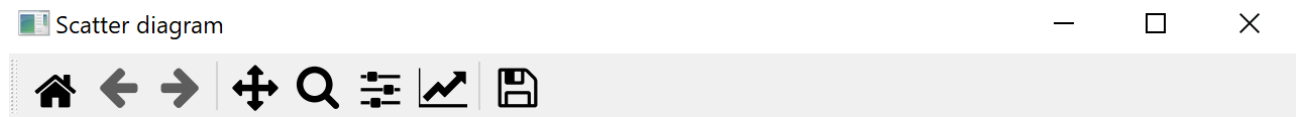
for x, y in arrx, arry:

$\text{sum} += ((x - \text{середнє по } x) / sx) * ((y - \text{середнє по } y) / sy)$

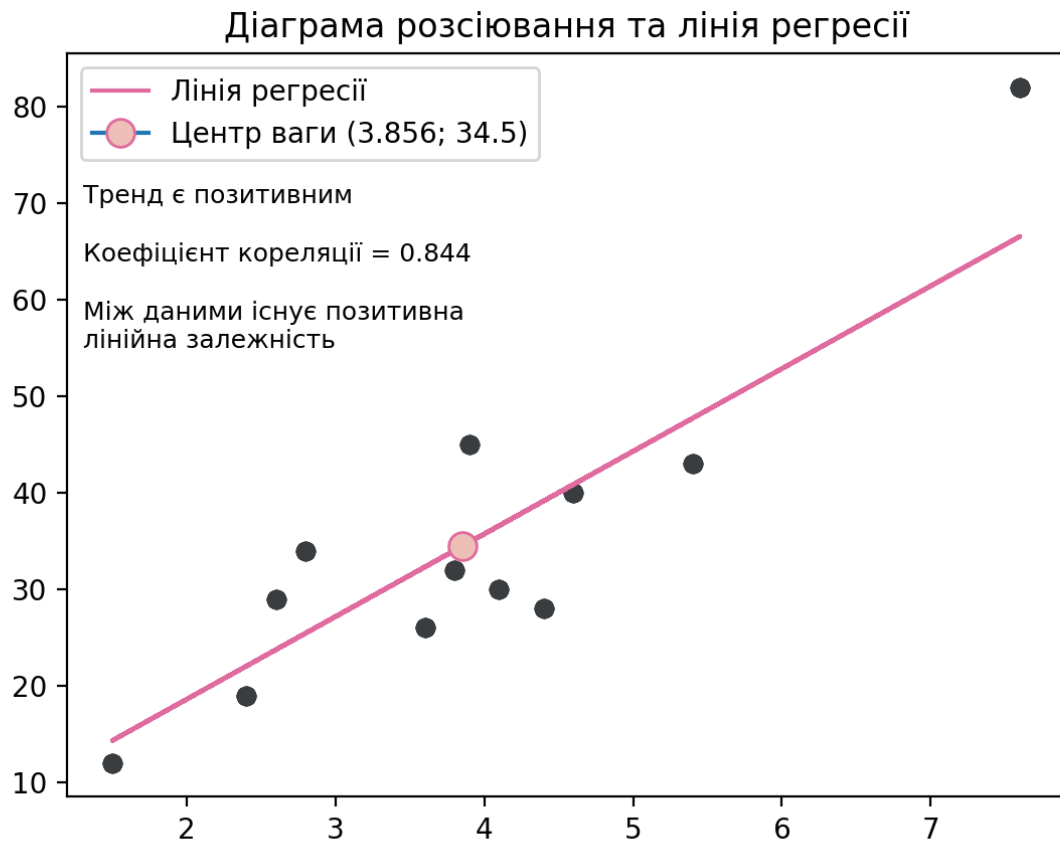
$r = \text{sum} / (\text{к-ть елементів} - 1)$

Випробування алгоритму

10 елементів:



100 елементів:



Висновок

Під час виконання лабораторної роботи було розроблено програму зчитування та аналізу даних з документу та запису вихідних даних. Було побудовано алгоритми знаходження коваріації, лінії регресії, коефіцієнта кореляції, визначено центр ваги та тренд. Побудовано діаграму розсіювання.