

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

Виконав:	Якубець М. В.	Перевірила:	Вечерковська А. С.
Група	ІПЗ-22	Дата перевірки	
Форма навчання	денна	Оцінка	
Спеціальність	121		
2022			

Назва: Двовимірна статистика.

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

Математична модель:

N - кількість елементів

$$\bar{x} = \left(\sum_{x \in X} (x) \right) / N$$

$$\text{cov}(x, y) = \frac{\sum_{i=1}^N (x_i * y_i)}{N} - \bar{x} * \bar{y} - \text{коваріація}$$

$$\text{var}(x) = \frac{\sum_{x \in X} f_i * (x_i - \bar{x})^2}{N}$$

$$b_0 = \frac{\text{cov}(x, y)}{\text{var}(x)}$$

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

$$y = b_0 * x + b_1 - \text{рівняння лінії регресії у від } x$$

$$\sigma = \sqrt{\text{var}(x)}$$

$$r = \frac{\text{cov}(x, y)}{\sigma_x * \sigma_y} - \text{коефіцієнт кореляції}$$

Псевдокод алгоритму:

```
from functools IMPORT reduce
from math IMPORT sqrt
```

```
DEFINE FUNCTION calculate_mean(list : list[float]) -> float:
```

```
    RETURN reduce(lambda x, y: x + y, list) / len(list)
```

```
DEFINE FUNCTION calculate_gravity_center(x_list: list[float], y_list: list[float]) -> tuple[float, float]:
```

```
    RETURN calculate_mean(x_list), calculate_mean(y_list)
```

```
DEFINE FUNCTION calculate_covariance(
```

```
    x_list: list[float], y_list: list[float],
```

```
gravity center: tuple[float, float]
) -> float:
    RETURN sum(
        map(lambda : [0] * [1], zip(x list, y list))
    ) / len(x list) - gravity center[0] * gravity center[1]
```

```
DEFINE FUNCTION calculate dispersion(list : list[float],
mean: float) -> float:
```

```
    RETURN reduce(lambda x, y: x + y, map(lambda x: (x -
mean) ** 2, list )) / len(list )
```

```
DEFINE FUNCTION calculate_average_square deviation(list :
list[float], mean: float) -> float:
```

```
    RETURN sqrt(calculate dispersion(list , mean))
```

Випробування алгоритму:

Задача №1:

Тренд позитивний

Задача №2:

Центр ваги = (35.0, 3.89)

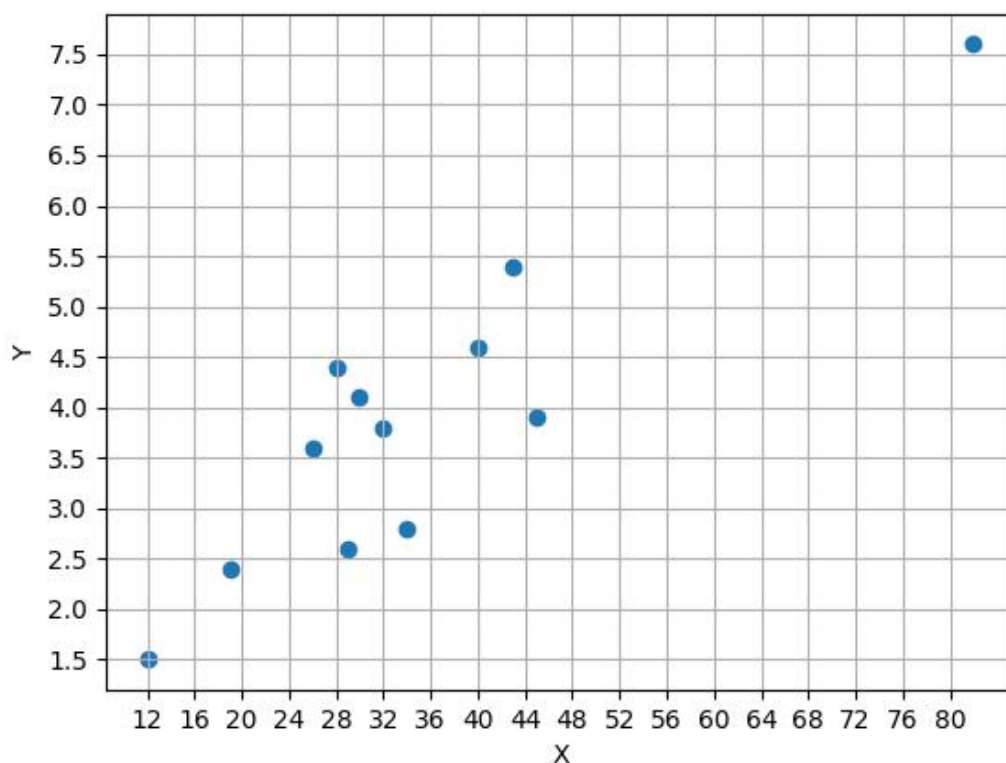
Коваріація = 23.0

Задача №3:

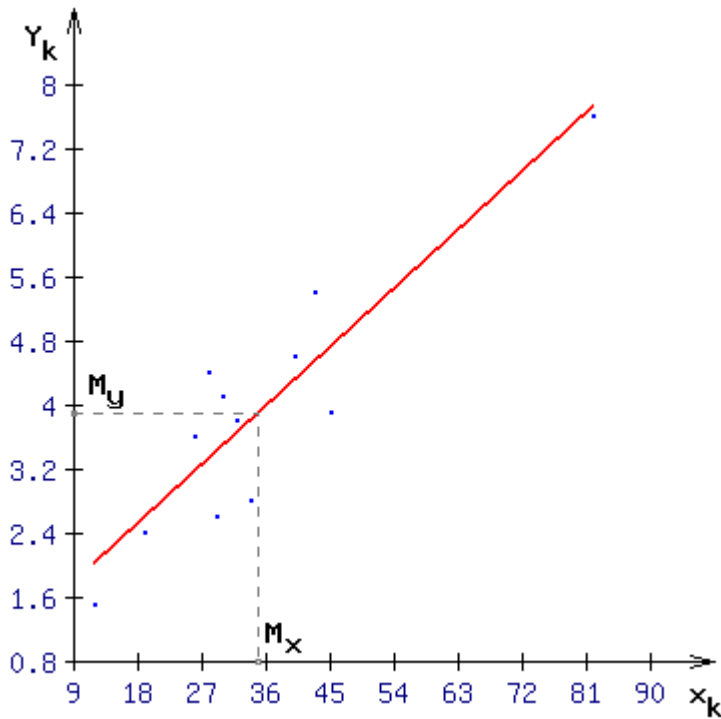
Рівняння лінії регресії: $y = 0.08x + 1.04$

Задача №4:

Коефіцієнт кореляції = 0.9



$$\text{cov}(X, Y) = 23.000000$$



$$R_{xy} = 0.901001$$

Висновок: Навчився використовувати на практиці набуті знання про міри в двовимірній статистиці. Намалював діаграму розсіювання для даних. Зрозумів залежність між трендом та коефіцієнтом кореляції. Набув навички розрахунку та використання коваріації. На практиці побачив зв'язок між коефіцієнт кореляції та діаграмою розсіювання для даних: чим ближче коефіцієнт до одиниці, тим “стрункіша” до прямої лінії точки.