

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:	Шевчук Максим Юрійович	Перевірів:	Вечерковська Анастасія Сергіївна
Група	ІПЗ-22	Дата перевірки	
Форма навчання	денна	Оцінка	
Спеціальність	121		
2022			

Тема: ДВОВИМІРНА СТАТИСТИКА

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

Постановка задачі: для розв'язання даної лабораторної роботи необхідно побудувати декілька математичних моделей відповідно до завдань.

Для побудови діаграми розсіювання необхідно застосувати бібліотеку Matplotlib

Математична модель:

Функція для побудови лінії розсіювання:

$$y = k + mx.$$

, де

$$m \equiv b_1 = \frac{\text{cov}(X, Y)}{\text{Var}(X)}$$

$$k \equiv b_0 = \bar{y} - b_1 \bar{x}.$$

Формула кореляції

$$r = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

Формула коваріації

$$\text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

Центр ваги

The point $G(\bar{x}, \bar{y})$ is called the center of gravity of the data.

Псевдокод алгоритму:

Коваріація:

```
data_x
data_y
average_x
average_y

covariation_result = 0
    for i in range data_sum.length
covariation_result += (data_x[i] - average_x)*(data_y[i] - average_x)
    covariation_result /= data_x.length
```

Кореляція:

```
data_x
data_y
average_x
average_y
deviation_x
deviation_y
covariation

for i in range data_sum.length
{
covariation_result += (data_x[i] - average_x)*(data_y[i] - average_x)
}
covariation_result /= data_x.length

corelation = covariation / sqrt(deviation_x *
deviation_y)
```

Коефіцієнт m лінії регресії:

```
data_x
```

```
deviation_x
```

```
m = covariation/deviation_x
```

Коефіцієнт k лінії регресії:

```
m_coefficient
```

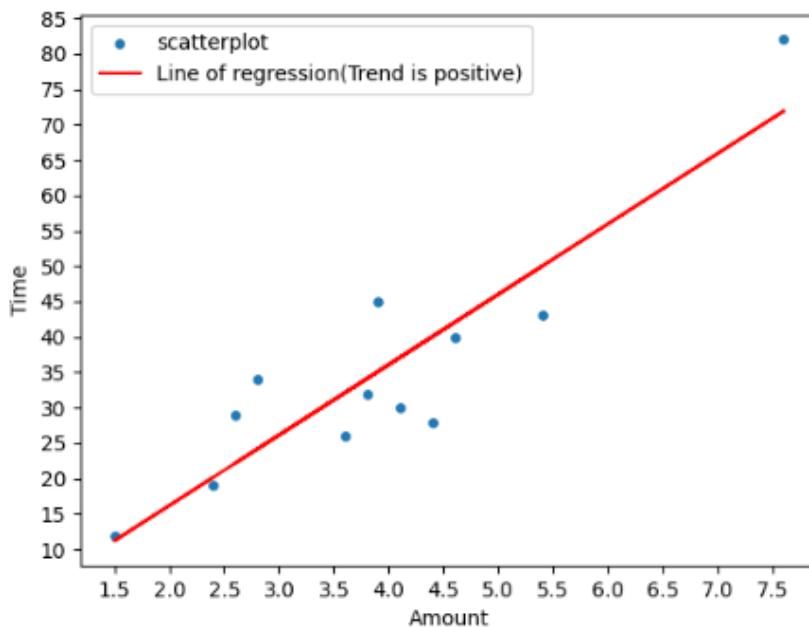
```
average_x
```

```
average_y:
```

```
k = average_y - average_x*m_coefficient
```

Випробування алгоритму:

```
Center of weight G(3.892;35.0)
Covariation is 23.0
Line of regression is  $y=9.953x + -3.735$ 
Coefficient of correlation is 0.901
```



Висновок: на цій лабораторній роботі було розглянуто різноманітні методи обробки масивів даних з точки зору двовимірної статистики. Було побудовано діаграму розсіювання та лінію регресії даних.