

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 № 2

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

Назва: Лінійне перетворення та Графічне зображення даних.

Мета: Навчитись використовувати на практиці набуті знання про лінійні перетворення та графічне зображення даних.

Постановка задачі:

1. Знайдіть Q_1 , Q_3 та P_{90} .
2. Знайдіть середнє та стандартне відхилення цих оцінок.
3. Через незадоволення низькими оцінками викладач вирішив використати шкалу форми $y = ax + b$, щоб відредагувати оцінки. Він хотів, щоб середнє значення масштабних оцінок становило 95, а оцінка 100, щоб залишалася рівною 100.
4. Показати дані за допомогою діаграми "стовбур – листя".
5. Відобразити дані за допомогою коробкового графіка.
6. Зробити висновок.

Математична модель:

N - кількість елементів

$$element = (0.25 * (N + 1))^{th}$$

$$Q_1 = element^{th} + 0.25 * (element^{th+1} - element^{th})$$

$$element = (0.75 * (N + 1))^{th}$$

$$Q_3 = element^{th} + 0.75 * (element^{th+1} - element^{th})$$

$$element = (\frac{k}{100} * (N + 1))^{th}$$

$$P_{90} = element^{th} + \frac{k}{100} * (element^{th+1} - element^{th}), \text{ де } k - \text{ той перцентиль, який треба знайти}$$

$$\bar{x} = (\sum_{x \in X} (x)) / N, \text{ де } N - \text{ кількість елементів}$$

$$s_x^2 = \frac{\sum_{x \in X} f_i \cdot (x_i - \bar{x})^2}{N - 1}$$

$$s_x = \sqrt{s_x^2(x)}$$

Середнє квадратичне відхилення розподілу =

$$\sigma = \sqrt{\frac{\sum_{x \in X} f_i \cdot (x_i - \bar{x})^2}{N}}$$

$$\text{Стандартне відхилення} = \frac{element - \bar{x}}{\sigma}$$

$y_1 = a * x_1 + b$, де $y_1 = 100$, $x_1 = 100$

$y_2 = a * x_2 + b$, де y_2 дорівнює тому числу, до якого треба прирівняти, а $x_2 = \bar{x}$

Шукаємо а та b. Потім для кожного елементу x робимо: новий елемент = $x * a + b$

Псевдокод алгоритму:

```
from functools IMPORT reduce
from math IMPORT sqrt
IMPORT numpy as np
```

```
DEFINE FUNCTION calculate_q or p(list : list[int],
fraction: float) -> float:
    SET index TO int(fraction * (len(list) + 1)) - 1
    RETURN list [index] + fraction * (list [index + 1] -
list [index])
```

```
DEFINE FUNCTION calculate mean(list : list[int]) -> float:
    RETURN reduce(lambda x, y: x + y, list) / len(list)
```

```
DEFINE FUNCTION _calculate_numerator dispersion(list :
list[int], mean: float) -> float:
    SET list TO map(lambda x: (x - mean) ** 2, list)
    RETURN reduce(lambda x, y: x + y, list)
```

```
DEFINE FUNCTION calculate_average square deviation(list :
list[int], mean: float) -> float:
    SET numerator TO calculate_numerator dispersion(list,
mean)
    RETURN sqrt(numerator / (len(list) - 1))
```

```
DEFINE FUNCTION calculate_standard deviation(list :
list[int], mean: float) -> float:
    SET numerator TO calculate_numerator dispersion(list,
mean)
    RETURN sqrt(numerator / len(list))
```

```
DEFINE FUNCTION calculate_z_score(integer: int, mean:
float, standard deviation: float) -> float:
    RETURN (integer - mean) / standard deviation
```

```
DEFINE FUNCTION get a and b(mean: float, to: int) ->
np.ndarray:
```

```

SET coefficients TO [[100, 1], [mean, 1]]
SET answers TO [100, to]
RETURN np.linalg.solve(coefficients, answers)

```

```

DEFINE FUNCTION get_rearranged_list_for_teacher(list :
list[int], mean: float) -> list[float]:
    SET a, b TO get a and b(mean, 95)
    RETURN [round(value * a + b, 2) FOR value IN list ]

```

```

DEFINE FUNCTION create_stems_with_empty_leafs(list :
list[str]) -> dict[str, list]:
    SET min TO int(list [0][:-1]) IF list [0][:-1] else 0
    SET max TO int(list [-1][:-1]) + 1
    RETURN {str(k): [] FOR k IN range(min , max )}

```

```

DEFINE FUNCTION create_stem_and_leaf_data(list :
list[int]) -> dict[str, list[str]]:
    SET list TO convert list items to type(str, list )
    SET stem_and_leaf_data TO
create_stems_with_empty_leafs(list )
    FOR IN list :
        SET stem, leaf TO [:-1], [-1]
        SET stem TO stem IF stem else "0"
        stem and leaf data[stem].append(leaf)
    RETURN stem and leaf data

```

Випробування алгоритму:

Задача №1:

$$Q_1 = 62.75$$

$$Q_3 = 93.75$$

$$P_{90} = 99.5$$

Задача №2:

Середнє квадратичне відхилення = 18.1

Стандартне відхилення = -1.99

Задача №3:

Відредаговані оцінки = [88.37, 92.64, 93.22, 93.41, 94.19, 94.19, 96.9, 98.06, 99.03, 100.0]

Задача №4:

Діаграма стовбур-листя

Stem Leaf

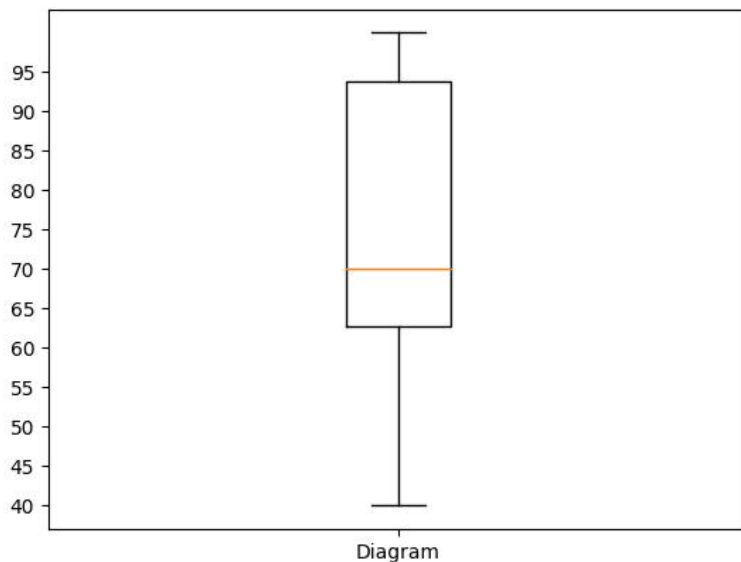
4 | 0

5 |

```

6 | 2 5 6
7 | 0 0
8 | 4
9 | 0 5
10 | 0

```



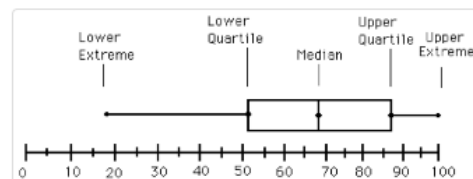
Quartile calculator Q1, Q3

For quartiles Q1, Q3 calculation, please enter numerical data separated with comma (or space, tab, semicolon, or newline). For example: -235.4 -303.8 838.9 271.2 903.7 269.6 596.4 285.8 632.0 383.9 508.2 144.6 769.6

40, 65, 62, 70, 100, 90, 66, 70, 95, 84

Recalculate

Reset



*For low count distributions, there is no universal agreement on selecting the quartile values (divide the ordered data set into two halves and then next halving...). If there are even number of data points, all methods give the same results.

Calculation:

Statistical file:

{40, 65, 62, 70, 100, 90, 66, 70, 95, 84}

Quartile Q1: 64.25

Quartile Q2: 70

Quartile Q3: 91.25

Other statistical characteristics:

Average (mean): $\mu=74.2$

Absolute deviation: 144.4

Mean deviation: 14.44

Minimum: 40

Maximum: 100

Variance: 294.96

Standard deviation $\sigma=17.174399552823$

Corrected sample standard deviation $s=18.103406677566$

Z-score: {-1.9913, -0.5357, -0.7104, -0.2446, 1.5022, 0.92, -0.4775, -0.2446, 1.2111, 0.5706}
Count items: 10

Висновок: Навчився використовувати на практиці набуті знання про лінійні перетворення та графічне зображення даних. Перевінив зв'язок між квантилями, перцентилями та модою. Отримав досвід побудови діаграми стовбур-листя та коробкового графіка. Виявив, що чим більше даних, тим більше квантили та перцентилю.