

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 № 1

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

Назва: Центральні тенденції та міра дисперсії.

Мета: Навчитись використовувати на практиці набуті знання про центральні тенденції та міри.

Постановка задачі:

1. Побудувати таблицю частот та сукупних частот для переглянутих фільмів.
Визначити фільм, який був переглянутий частіше за інші.
2. Знайти Моду та Медіану даної вибірки.
3. Порахувати Дисперсію та Середнє квадратичне відхилення розподілу.
4. Побудувати гістограму частот для даного розподілу.
5. Зробити висновок з вигляду гістограми, про закон розподілу.

Математична модель:

Mean Median Mode Formula



$$\text{Mean} = \frac{\text{Sum of Observations}}{\text{Total Number of Observations}}$$

$$\text{If 'n' is odd: Median} = \left(\frac{n+1}{2} \right)^{\text{th}} \text{ term}$$

$$\text{If 'n' is even: Median} = \frac{\left(\frac{n}{2} \right)^{\text{th}} \text{ term} + \left(\frac{n}{2} + 1 \right)^{\text{th}} \text{ term}}{2}$$

$$\text{Mode} = L + h \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)}$$

$$D = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Дисперсію = _____

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}.$$

Середнє квадратичне відхилення розподілу =

Псевдокод алгоритму:

```

1  from functools import reduce
2  from math import sqrt
3  from typing import TextIO
4
5  from matplotlib import pyplot as plt
6
7
8  def task1(movies_views: list[int], file: TextIO) -> None:
9      row_format = "{:>3} {:>10}"
10     file.write("Таблиця частот:\n")
11     file.write("Фільм   Перегляди\n")
12     for index, movie_views in enumerate(movies_views, 1):
13         file.write(f"{row_format.format(index, movie_views)}\n")
14     file.write("-----\n")
15
16     checked_movies = []
17     file.write("Таблиця сукупних частот:\n")
18     file.write("Частота   Перегляди\n")
19     for movie_views in movies_views:
20         if movie_views in checked_movies:
21             continue
22         frequency = movies_views.count(movie_views)
23         checked_movies.append(movie_views)
24         file.write(f"{row_format.format(frequency, movie_views)}\n")
25
26     file.write("-----\n")
27
28     max_views = max(movies_views)
29     for index, movie_views in enumerate(movies_views, 1):
30         if max_views == movie_views:
31             file.write(f"Фільм, переглянутий найчастіше: {index} -- {max_views}\n")
32
33  def task2(movies_views: list[int], file: TextIO) -> None:
34     file.write(f"Мода: {max(movies_views)}\n")
35
36     movies_views = sorted(movies_views)
37     half_length = len(movies_views) / 2
38     file.write(f"Медіана: {(movies_views[int(half_length - 0.5)] + movies_views[int(half_length)])} / 2\n")
39
40
41  def task3(movies_views: list[int], file: TextIO) -> None:
42     movies_views_length = len(movies_views)
43     mean = reduce(lambda x, y: x + y, movies_views) / movies_views_length
44     result = 0
45     for movie_views in movies_views:
46         result += (movie_views - mean) ** 2
47     file.write(f"Дисперсія: {result / (movies_views_length - 1)}\n")
48

```

```

49     file.write(f"Середнє квадратичне відхилення розподілу: {sqrt(result / movies_views_length)}\n")
50
51
52     def task4(movies_views: list[int]) -> None:
53         plt.figure(figsize=(15, 6), dpi=80)
54         plt.xlabel("Movies")
55         plt.ylabel("Views")
56         plt.bar([f"{_ * ' '}" for _ in range(len(movies_views))], movies_views)
57         plt.savefig("result.jpg")
58
59
60     def main() -> None:
61         file_name = input("Enter file name: ")
62         try:
63             with open(file_name) as file:
64                 movies_views = list(map(int, file.read().splitlines()[1:]))
65         except FileNotFoundError:
66             print("Напишіть дійсну назву файлу...")
67             return
68
69         with open("result.txt", "w", encoding="utf-8-sig") as file:
70             file.write("Завдання №1:\n")
71             task1(movies_views, file)
72
73             file.write("Завдання №2:\n")
74             task2(movies_views, file)
75
76             file.write("Завдання №3:\n")
77             task3(movies_views, file)
78
79             task4(movies_views)
80
81
82     if __name__ == "__main__":
83         main()
84

```

Випробування алгоритму:

Завдання №1:

Таблиця частот:

Фільм Перегляди

1	1
2	66
3	75
4	1
5	1
6	12
7	10
8	97
9	12
10	66

Таблиця сукупних частот:

Частота Перегляди

3	1
2	66
1	75
2	12
1	10
1	97

Фільм, переглянутий найчастіше: 8 -- 97

Завдання №2:

Мода: 97

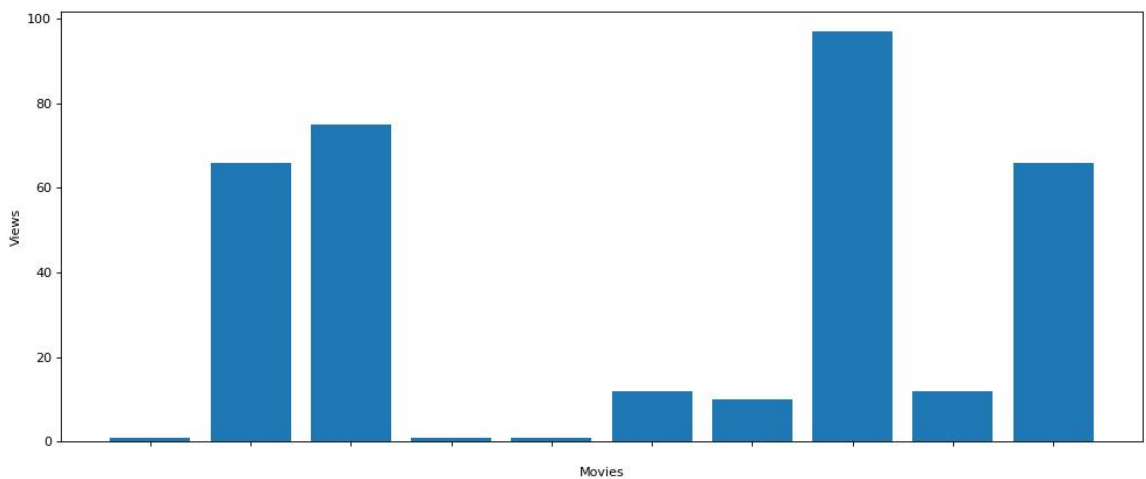
Медіана: 12.0

Завдання №3:

Дисперсія: 1389.877777777778

Середнє квадратичне відхилення розподілу: 35.3679233204326

Завдання №4:



Медіана набору чисел

1 66 75 1 1 12 10 97 12 66

Разделитель

пробел

Вычислить

1 66 75 1 1 12 10 97 12 66

Медіана = 12

Дисперсія (s^2)

1389.88

Стандартное отклонение (s)	37.2811
Считать (n)	10
Сумма (Σx)	341
Иметь в виду (\bar{x})	34.1
Коэффициент дисперсии	1.0933
Сумма площадей (SS)	12508.9

Дисперсія =

Середнє квадратичне відхилення розподілу =

Standard Deviation Calculator

Result

Standard Deviation, σ : 35.367923320433

Count, N: 10
Sum, Σx : 341
Mean, μ : 34.1
Variance, σ^2 : 1250.89

Steps

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}.$$

$$\begin{aligned}\sigma^2 &= \frac{\Sigma(x_i - \mu)^2}{N} \\ &= \frac{(1 - 34.1)^2 + \dots + (66 - 34.1)^2}{10} \\ &= \frac{12508.9}{10} \\ &= 1250.89 \\ \sigma &= \sqrt{1250.89} \\ &= 35.367923320433\end{aligned}$$

Висновок: Навчився використовувати на практиці набуті знання про центральні тенденції та міри. Побачив та довів зв'язок моди, медіани, дисперсії та середньоквадратичного відхилення розподілу з початковими даними. Отримав досвід роботи з гістограмою.