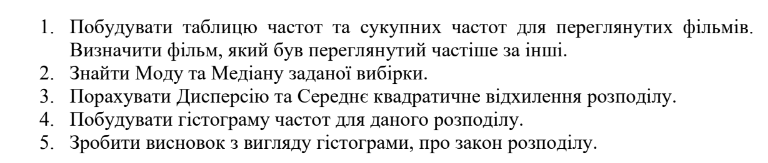
|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**  Дисципліна  **«Ймовірнісні основи програмної інженерії»**  **Лабораторна робота № 1**  **«Центральні тенденції та міра дисперсії»** | | | |
| **Виконав:** | Зганяйко В.В. | **Перевірила**: | Вечерковська А. |
| Група | ІПЗ-21(1 підгрупа) | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

**Тема:** центральні тенденції та міра дисперсії.

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

**Завдання:**



**Постановка задачі:**

Перш за все, аналіз даного завдання необхідно почати із вхідних даних та яким чином ми їх будемо отримувати. За умовою завдання сказано, що вхідні дані програма буде отримувати із текстового файлу, який складається з М+1 рядків з числами, де перше значення – це кількість елементів у файлі, а всі інші – кількість переглядів n-го фільму. Отже, спочатку необхідно буде зчитати дані з файлу у певний масив. Після цього всі дії будуть проходити над даним масивом та його даними. Розраховуємо усі необхідні нам значення за певними формулами. Далі необхідно усі ці дані записати також у окремий файл, щоб наша програма працювала як чорний ящик.

**Математична модель:**

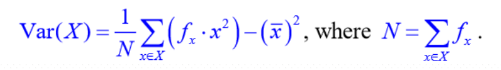
**Завдання 1**

За умовою даного завдання необхідно побудувати таблицю частот та сукупних частот переглянутих фільмів. Також визначити фільм з найбільшою кількістю переглядів. Для побудови таблиці частот необхідно обчислити відносну частоту для кожного фільму, для цього ми беремо суму переглядів усіх фільмів, а потім ділимо кількість переглядів фільму на ось цю суму та множимо на 100. Далі необхідно обрахувати сукупну частоту для кожного фільму, вона також обраховується за певним принципом: ми беремо суму переглядів попередніх фільмів та додаємо до неї кількість переглядів даного фільму. Щоб визначити фільм з найбільшою кількістю переглядів, необхідно просто визначити найбільше значення масиву, в якому зберігаються перегляди.

**Завдання 2**

За умовою даного завдання необхідно знайти моду та медіану вибірки. Мода вибірки – це значення, яке повторюється найчастіше, в нашому випадку це буде індекс максимального елементу масиву, що містить в собі частоти. Медіана – це середній елемент вибірки. Якщо даний масив матиме непарну довжину, то медіаною буде середній елемент вибірки, якщо ж парну, то медіана – пів сума двох середніх елементів вибірки.

**Завдання 3**

За умовою даного завдання необхідно знайти дисперсію та середнє квадратичне відхилення вибірки. Щоб знайти дисперсію, необхідно спочатку обрахувати середнє значення вибірки, та підставити його у певну формулу: 

Для знаходження відхилення також є певна формула:



**Завдання 4**

Для побудови гістограми частот даної вибірки, використаємо бібліотеку matplotlib та її метод hist.

**Завдання 5**

Закон розподілу – це поняття, яке показує множину можливих подій з ймовірністю їх настання. Необхідно проаналізувати гістограму та зробити певні висновки з її вигляду.

**Псевдокод алгоритму:**

**Отримання вхідних даних з файлу**

READ file with elements

CREATE array

WHILE element in file != NULL

ADD element to array

CREATE new array2

ADD elements from array to array2 without the first one

**Завдання 1**

CALCULATE array2 sum

WHILE element in array2 != NULL

CALCULATE cumulative frequency for every element

CALCULATE frequency for every element

WRITE INTO file all information we need

FIND max element in array2

WRITE INTO file max element

**Завдання 2**

FIND index of max element

WRITE INTO file index of max element

CREATE new array with all elements

IF length of new array % 2 == 1

mediana = newarray[newarray.Length/2]

ELSE

mediana = (newarray[newarray.Length/2] + newarray[newarray.Length/2-1]) / 2

WRITE INTO file mediana

**Завдання 3**

CALCULATE array2 sum

CALCULATE sum of all elements

CALCULATE average element

WRITE INTO file average element

CALCULATE dispersion(sum/sumoffrequences – average^2, where sum = sum(array2[i]\*(i+1)^2)

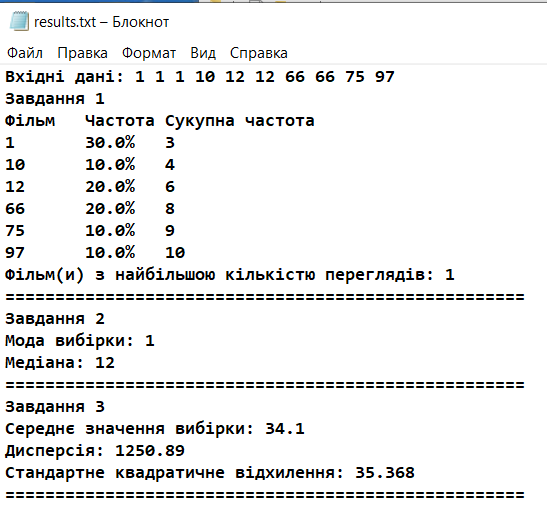
WRITE INTO file dispersion

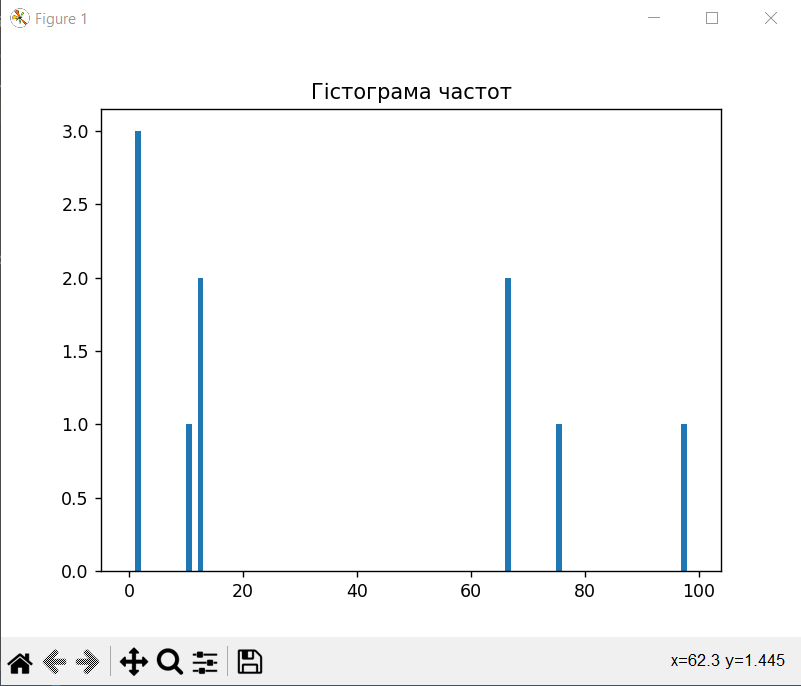
CALCULATE deviation (sqrt(dispersion))

WRITE INTO file deviation

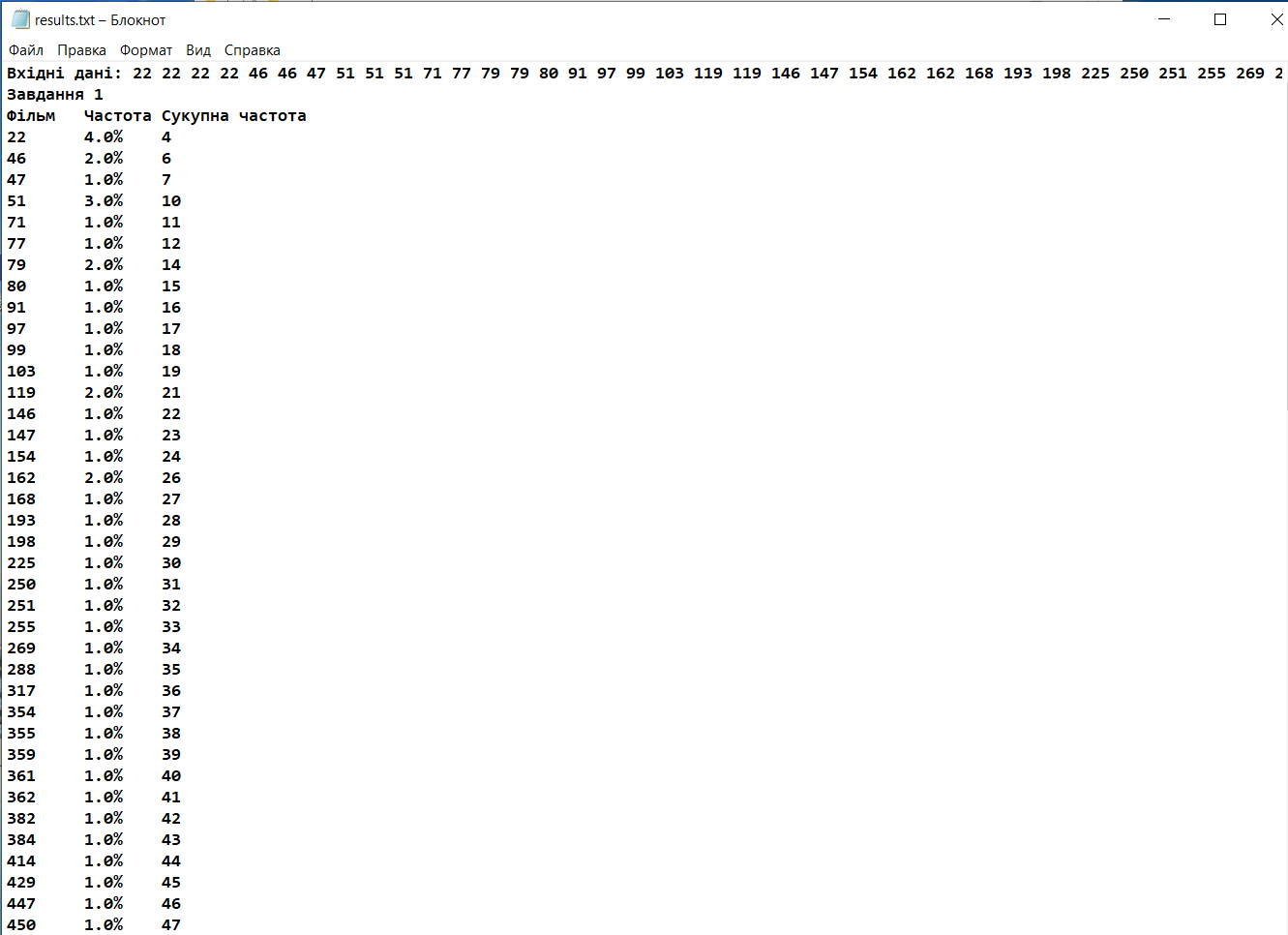
**Випробування алгоритму**:

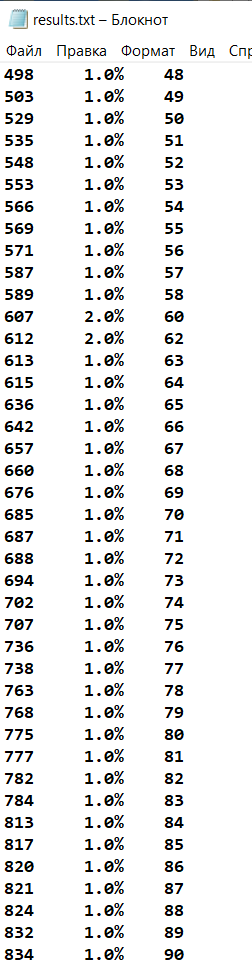
**Вихідні дані з 10 елементів:**

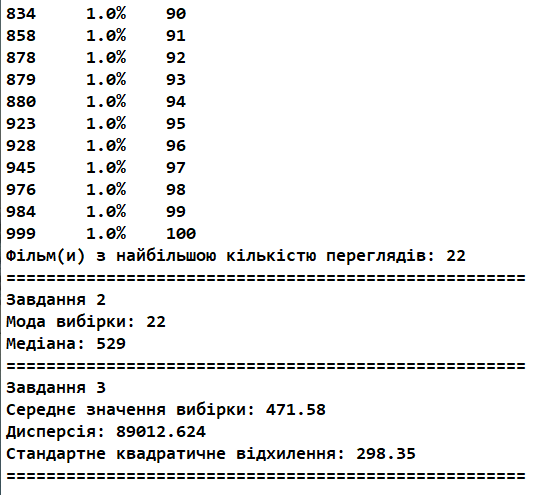


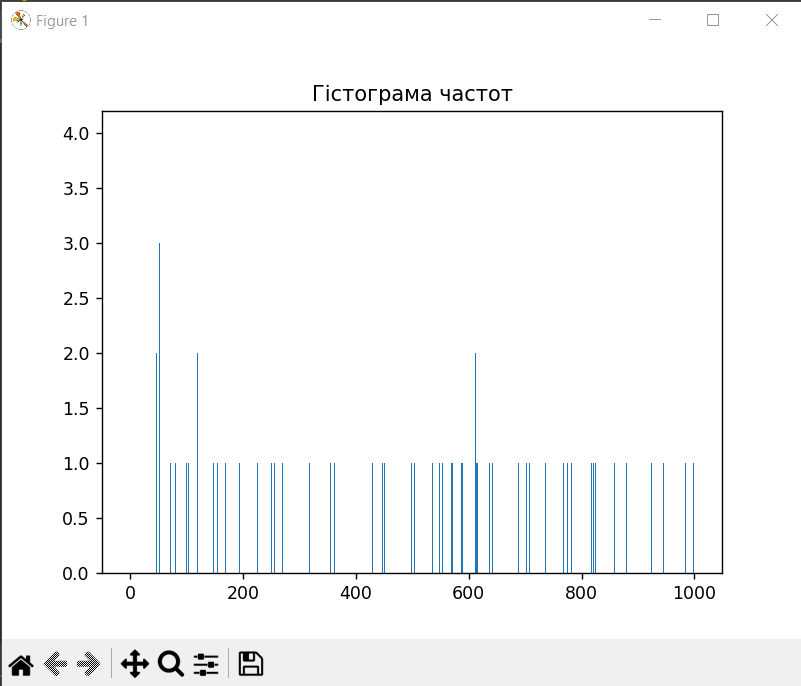


**Вихідні дані з 100 елементів:**









**Висновок:** У ході виконання даної лабораторної роботи, було виконано її завдання, а саме була розроблена програма, яка отримує назву файлу як вхідні дані та обраховує все необхідне, а потім записує результати до окремого файлу. Опановано використання на практиці навичок та знань про центральні тенденції та міри.