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»

Виконала:	Саніна В.О.	Перевірив:	Марцафей А.С.
Група	ІПЗ-21	Дата перевірки	
Форма навчання	денна	Оцінка	
Спеціальність	121		

2022

<u>Тема</u>: центральні тенденції та міра дисперсії.

<u>Мета</u>: навчитись використовувати на практиці набуті знання про центральні тенденції та міри.

Завдання:

- 1. Побудувати таблицю частот та сукупних частот для переглянутих фільмів. Визначити фільм, який був переглянутий частіше за інші.
- 2. Знайти Моду та Медіану заданої 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:

<u>Кумулятивна частота заданого значення змінної X_i </u> - сума абсолютних частот f усіх значень, менших або рівних їй — $F_i = f_1 + f_2 + f_3 + \dots$ F

 ${\underline{{\rm Moдa}}}$ — значення величини, що трапляється найчастіше в сукупності спостережень.

Медіана — серединне значення упорядкованої вибірки:

При непарній кількості елементів, медіана дорівнює сума_елементів/2, з округленням в більшу сторону.

При парній кількості елементів, медіана дорівнює сумі двох елементів: сума_елементів/2, з округленням в більшу сторону, та наступний після нього.

Дисперсія:
$$Var(X) = \frac{1}{N} * \sum_{x \in X} (f_x * x^2) - (\overline{x})^2$$
, де $N = \sum_{x \in X} f_x$

Середнє квадратичне відхилення розподілу: корінь квадратний від значення

дисперсії:
$$\sqrt{\frac{1}{N} * \sum_{x \in X(} f_x * x^2) - (\overline{x})^2}$$

Псевдокод алгоритмів:

moda = 0

for i in range(n-1): if arrdata[i]>max1: max1=arrdata[i]

```
Пошук сукупних частот:

Cfrequency = []

for i in range(n-1):

    if i == 0:

        x = arrdata[i]

        Cfrequency.append(x)
    else:

        x = Cfrequency[i-1]+arrdata[i]

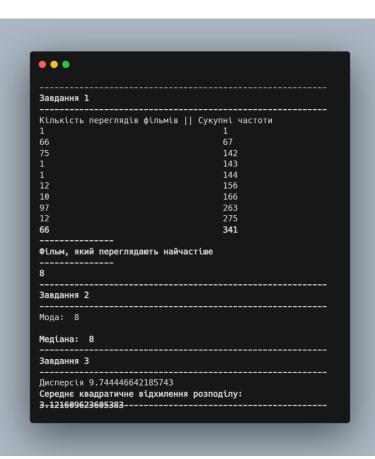
        Cfrequency.append(x)

Знаходження фільму, який переглядають частіше за інші:
    for i in range(n-1):
    if arrdata[i]>max1:
        max1=arrdata[i]

Знахолдження моди:
```

```
moda = i
Знаходження медіани:
if (sum1 \% 2 == 0):
  el1=sum1/2
  e12=e11+1
  res = (el1 + el2)/2
else: res = sum 1/2
for i in range(n-1):
 if ((res>sum2)) & (res<(sum2+arrdata[i])):
    print("Медіана: ", i+1)
    o.write("Файл на 10 елементів, медіана: " + str(i+1) + \n') # Записати рядок
    break
 sum2=sum2+arrdata[i]
Знаходження дисперсії та середнього квадратичного відхилення
розподілу:
for i in range(n-1):
 mean += (i+1)*arrdata[i]
  denominator+=arrdata[i]
 numeral += \operatorname{arrdata}[i]*(i+1)*(i+1)
disp=numeral/denominator
mean /=denominator
mean *= mean
disp -= mean
Створення гістограми на python:
fig, image = plt.subplots()
imageFilm = []
for i in range(n-1):
  imageFilm.append(i+1)
Imagefrequency = []
for i in range(n-1):
  Imagefrequency.append(arrdata[i])
bar colors = ['tab:blue']
image.bar(imageFilm, Imagefrequency, color=bar_colors)
image.set_ylabel('Frequency')
image.set_xlabel('Movie number')
image.set_title('histogram of frequencies')
```

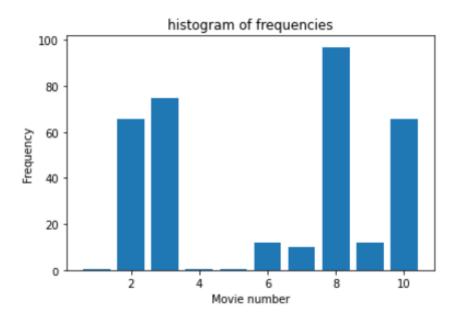
Демонстрація роботи алгоритмів на масиві з 10,100 та 1000 частот:

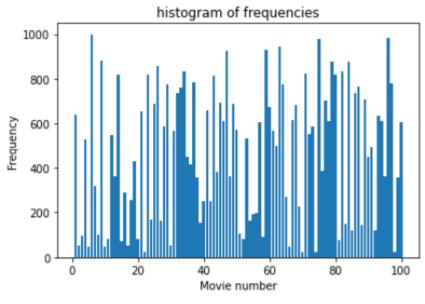


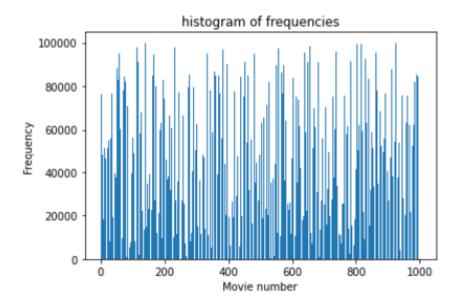
Фільм, який переглядають найчастіше
6
Завдання 2
 Мода: 6
Медіана: 54
Завдання З

•	
	 льм, який переглядають найчастіше
92	 !5
38	
Mc	
Me	едіана: 492
3	ивдання З
Ce	

Побудова гістограм для масивів на 10, 100 та 1000 частот:







<u>Висновок</u>: Під час виконання лабораторної, було розроблено алгоритми на теми центральних тенденцій та міри. Було побудовано таблицю частот та сукупних частот за даними зі вхідного файлу, знайдено моду та медіану заданої вибірки, пораховано дисперсію та середнє квадратичне відхилення розподілу, побудовано гістограму для заданої вибірки. За видом діаграми можемо сказати, що вона не симетрична, не має вираженого, діаграма не має виражених скосів.