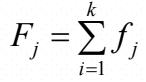
|  |  |  |  |
| --- | --- | --- | --- |
| МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені Тараса Шевченка  ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ  **Кафедра програмних систем і технологій**    Дисципліна  **«Ймовірнісні основи програмної інженерії»**  **Лабораторна робота № 1**  **Центральні тенденції та міра дисперсії** | | | |
| **Виконав:** | Баран В.М. | **Перевірила**: | Вечерковська А. С. |
| Група | ІПЗ-25мс | Дата перевірки | 14.11.2022 |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

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

**Хiд роботи**

**Завдання 1** Побудувати таблицю частот та сукупних частот для переглянутих фільмів. Визначити фільм, який був переглянутий частіше за інші.

1. Визначаємо частоту переглянутих фiльмiв та сукупну частоту за формулою:



2. Будуємо таблицю частот та сукупних частот, виводимо її на екран та записуємо до файлу.

3.Знаходимо найбiльш переглянутий фiльм по його частотi, виводимо його на екран та записуємо до файлу.

**Код:**

import matplotlib.pyplot as plt

import numpy as np

import math

my\_str\_2 = r'C:\Users\babam\Pictures\\'

f = open(r'C:\Users\babam\Documents\text.txt','w')

print('Введіть назву файла ')

name\_file=input()

with open(my\_str\_2+name\_file+".txt") as file:

    array = [row.strip() for row in file]

np\_array=np.unique(array)

print("\n")

count1=len(array)

if count1 > 12:

    np\_array = [int(x) for x in np\_array]

    array = [int(x) for x in array]

print("\nЗначення    ","Частота   ","Сукупна частота")

f.write("Значення    Частота   Сукупна частота\n")

ab = 0

for i in np\_array:

    ae = array.count(i)

    if(int(i)==0):

        ab=ae

    if(int(i)>0):

        ab=int(ab)+int(ae)

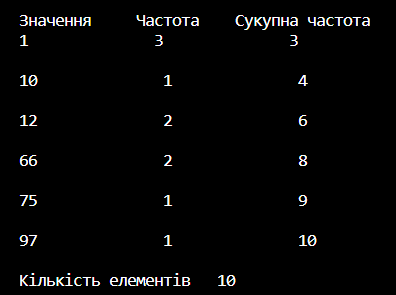
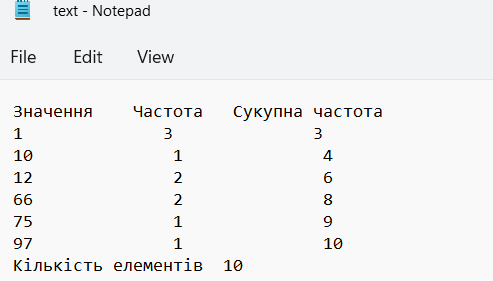
    f.write(str(i)+"              "+str(ae)+"              "+str(ab)+"\n")

    print(str(i)+"              "+str(ae)+"              "+str(ab)+"\n")

f.write("Кількість елементів  "+str(count1)+"\n")

print("Кількість елементів  ",count1)

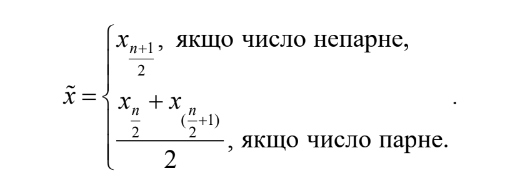
**Результат:**



**Завдання 2** Знайти Моду та Медіану заданої вибірки.

1. Знаходимо медiану з поточноi вибiрки за допомогою формули,

результат заносимо до документу та виводимо на екран.



1. Знаходимо моду та заносимо результат до документу та виводимо на екран.

**Код:**

most\_common = []

qty\_most\_common = 0

for item in np\_array:

    qty = array.count(item)

    if qty > qty\_most\_common:

        qty\_most\_common = qty

        most\_common = item

f.write("Мода  "+str(most\_common)+"\n")

print("Мода  ",most\_common)

sort = sorted(array)

for a in sort:

    nu = len(sort)/2

    n = int(sort[int(nu)-1])+int(sort[int(nu)])

f.write("Медіана  "+str(n/2)+"\n")

print("Медіана  ",n/2)

**Результат:**



**Завдання 3** Порахувати Дисперсію та Середнє квадратичне відхилення розподілу.

**Код:**

summ=0

for num in array:

    summ+=int(num)

avg = summ / len(array)

var = sum((float(x)-avg)\*\*2 for x in array) / len(array)

print("Дисперсія  ", var)

f.write("Дисперсія  "+str(var)+"\n")

print("Середнє квадратичне відхилення  ",math.sqrt(var))

f.write("Середнє квадратичне відхилення  "+str(math.sqrt(var))+"\n")

**Результат:**



**Завдання 4-5** Побудувати гістограму частот для даного розподілу та зробити висновок з вигляду гістограми, про закон розподілу.

Будуємо гiстограму частот для вибiрки з файлу input\_100.txt

**Код:**

if (count1>12):

    x=[]

    for i in range(19):

        x.append(22)

    for i in range(10):

        x.append(119)

    for i in range(6):

        x.append(216)

    for i in range(8):

        x.append(313)

    for i in range(6):

        x.append(410)

    for i in range(11):

        x.append(507)

    for i in range(14):

        x.append(604)

    for i in range(9):

        x.append(701)

    for i in range(11):

        x.append(798)

    for i in range(5):

        x.append(895)

    for i in range(1):

        x.append(999)

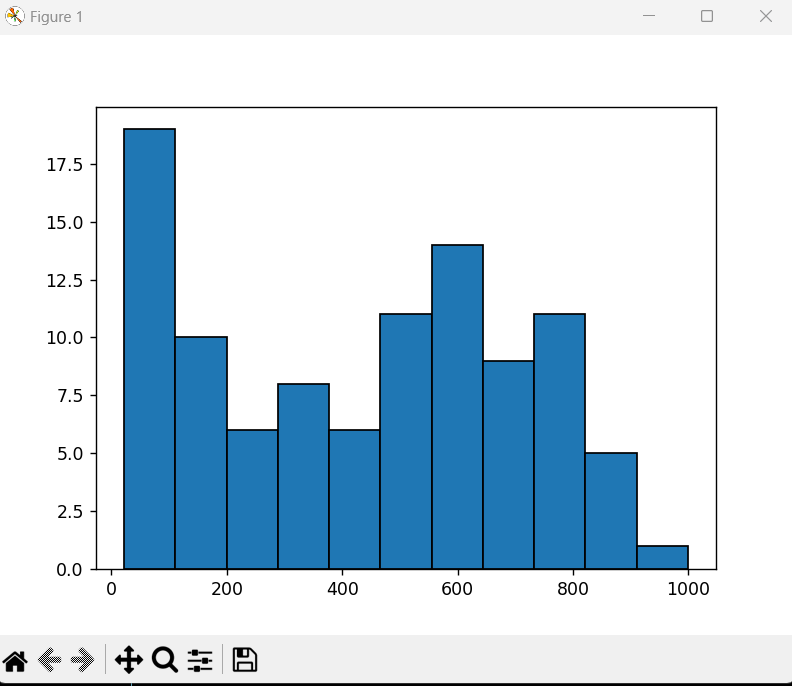
    m=m+1

n, bin, patches = plt.hist(x,bins=int(m),edgecolor='black')

plt.show()

file.close()

**Результат:**



**Висновок:** Гістограма скошена вправо. У випадку правого перекосу розподілу спостерігаємо: дані мають хвіст праворуч, середнє, медіана та мода розташовані в такій послідовності: мода < медіана < середнє.В цьому випадку клас [911, 100] є викидом, оскільки він має лише кілька спостережень.