|  |  |  |  |
| --- | --- | --- | --- |
| 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 № 3**  **«Двовимірна статистика»** | | | |
| **Виконала:** | Травіна Анастасія Ігорівна | **Перевірила**: | Марцафей А.С. |
| Група | ІПЗ-21 | Дата перевірки |  |
| Форма навчання | денна | Оцінка |  |
| Спеціальність | 121 |
| 2022 | | | |

**Мета**

Навчитись використовувати на практиці набуті знання про міри в двовимірній статистиці.

**Завдання**

Треба проаналізувати вхідний файл, що представляє собою текстовий файл із М+1 рядків. Перший рядок містить число М, де M — кількість значень у кожному з стовбців. Дослідження зроблено щодо суми покупки( стовбець 1) й часу, проведеного в супермаркеті(стовбець 2).Результати записуватимуться в окремий файл.

1. Намалюйте діаграму розсіювання для даних. Укажіть, чи існує тренд у даних.

Якщо так, то вкажіть, чи є це негативним трендом, чи позитивним.

2. Знайдіть центр ваги і коваріацію.

3. Знайти рівняння лініїї регресії y від x.

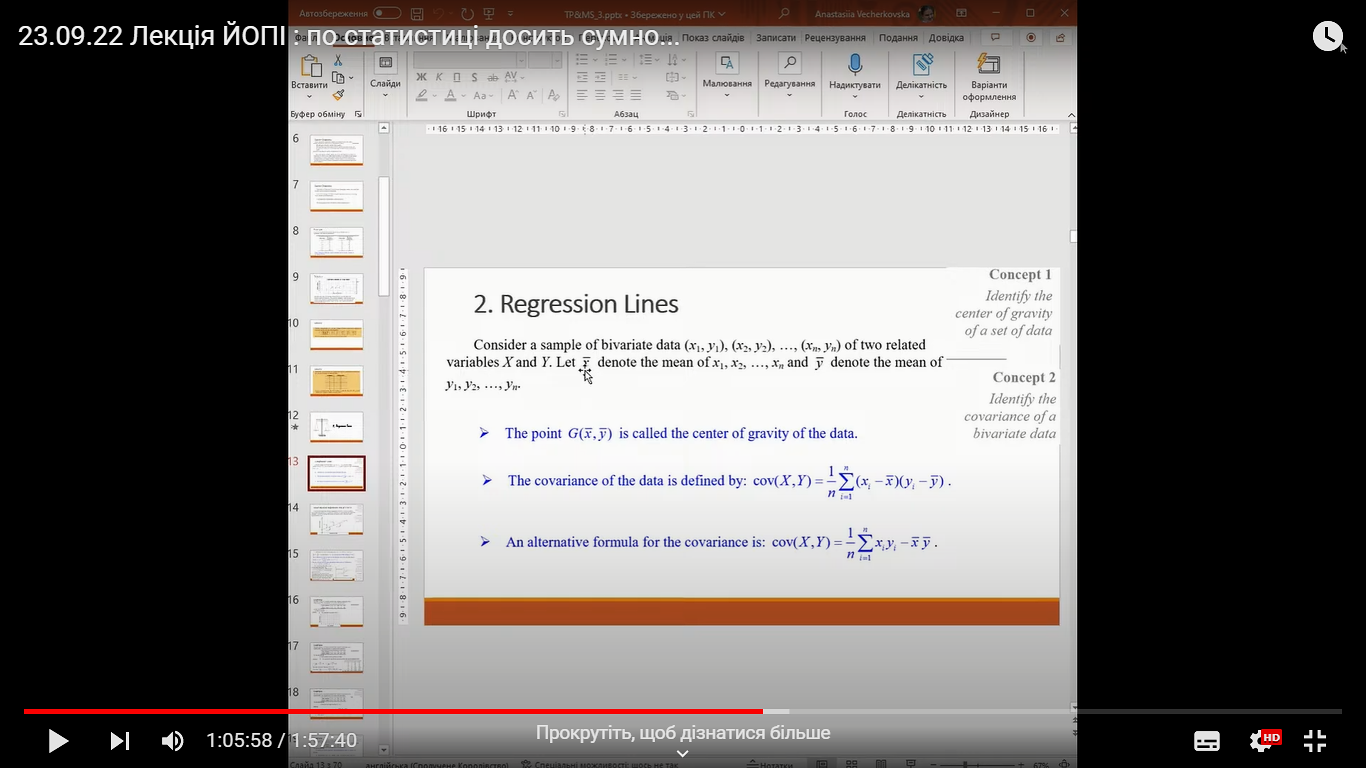
4. Розрахуйте коефіцієнт кореляції між даними.

5. Зробити висновок про залежності.

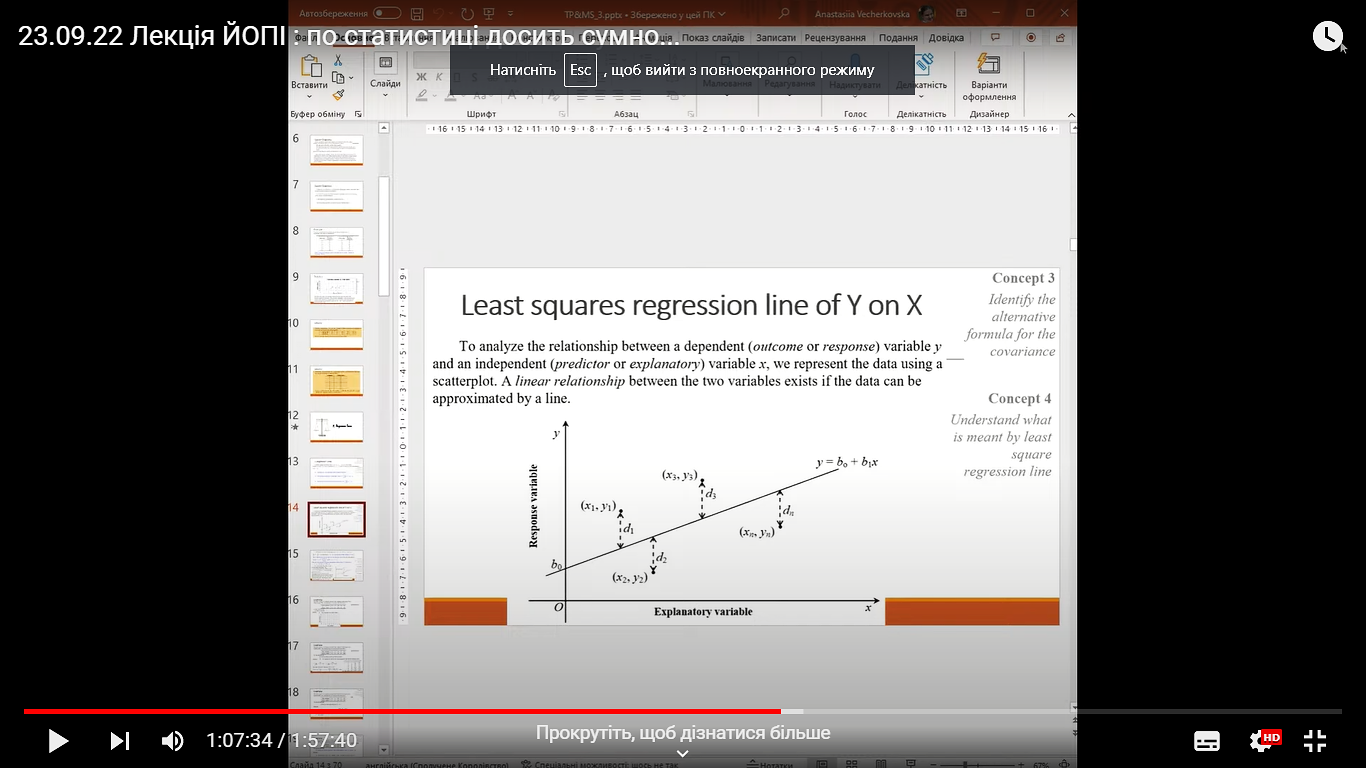
**Математична модель**

У даній лабораторній роботі було використано наступні формули:

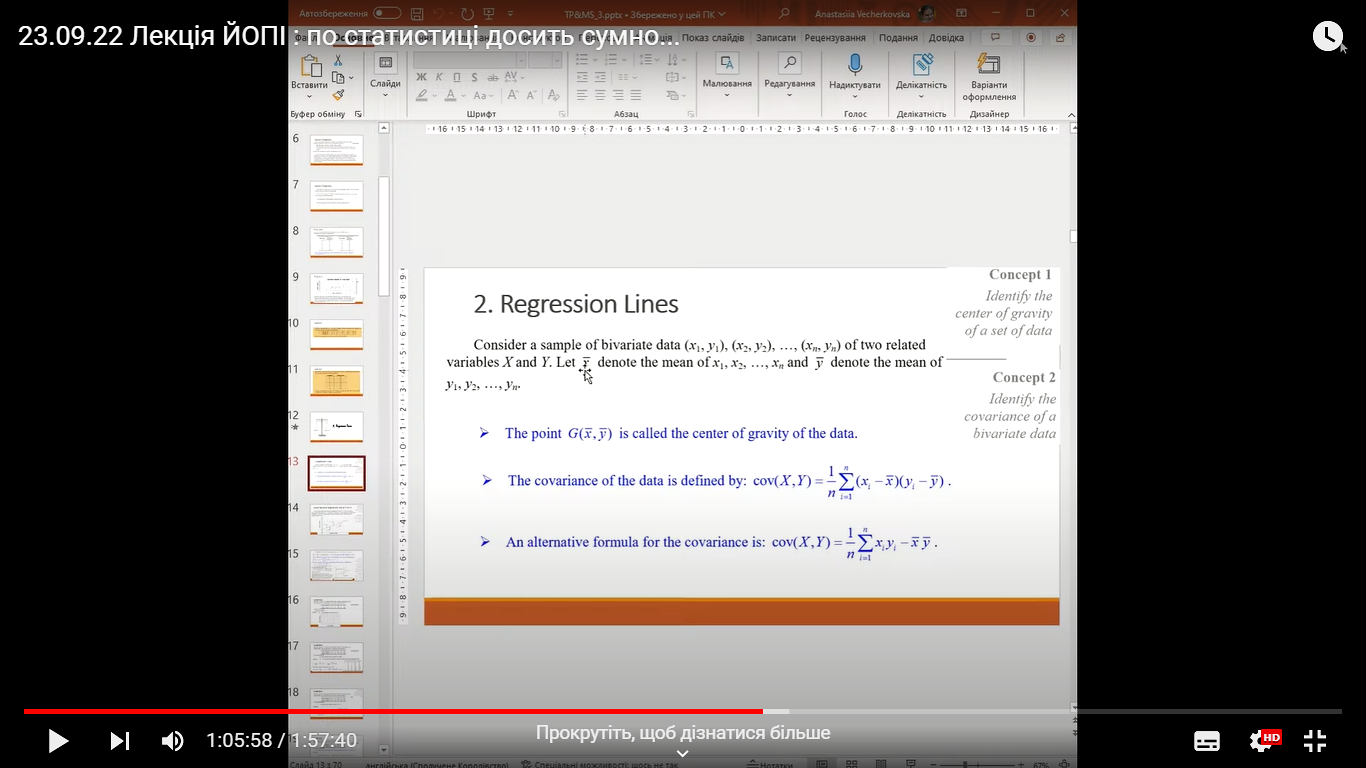
* Формула центра ваги



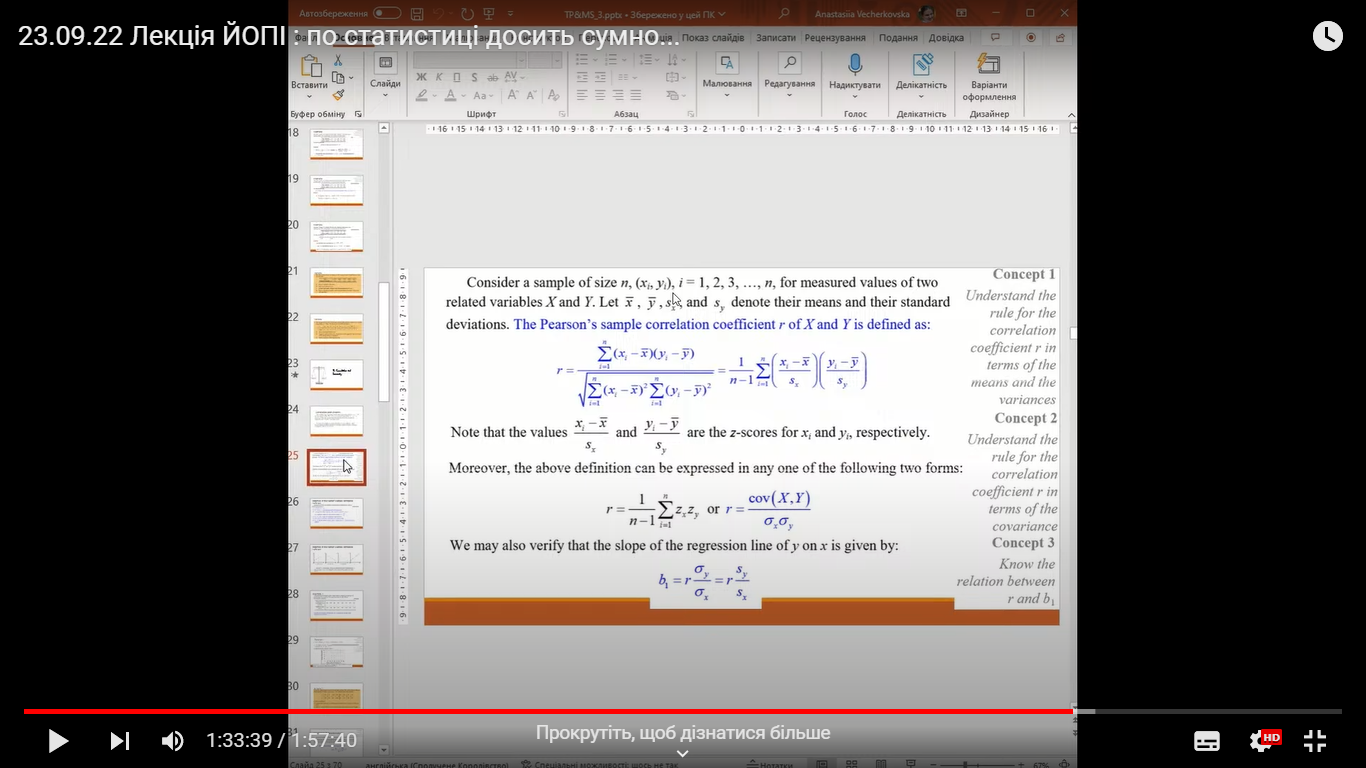
* Рівняння лінії регресії у від х



* Формула коваріації



* Формула коефіціента кореляції



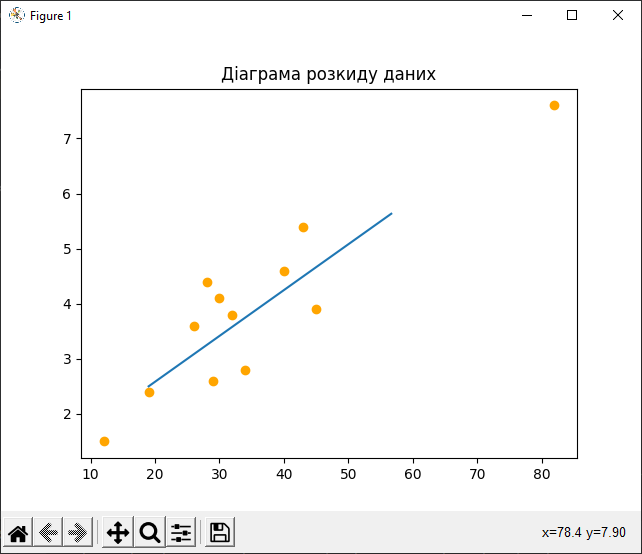
**Хід роботи**

Завдання 1

Код:

#Task 1  
new\_time, new\_paid =zip(\*[(time,paid) for time,paid in sorted(zip(time,paid))])  
for i in range(0,3):  
 beginp += new\_paid[i]  
 begint += new\_time[i]  
for i in range(paid.size-3,paid.size):  
 endp+=new\_paid[i]  
 endt += new\_time[i]  
A=[begint/3,beginp/3]  
B=[endt/3,endp/3]  
coords=list(zip(A,B))  
plt.scatter(time,paid,color='orange')  
plt.plot(coords[0],coords[1])  
plt.title('Діаграма розкиду даних')  
plt.show()  
temp=0  
temp2=0  
check=0  
new\_time, new\_paid =zip(\*[(time,paid) for time,paid in sorted(zip(time,paid))])  
file.write("Task 1\n")  
for i in range(0,paid.size-1):  
 check+=new\_paid[i+1]-new\_paid[i]  
if(check>0):  
 file.write("Trend is positive\n\n")  
else:  
 file.write("Trend is negative\n\n")

Результат (для файла input\_10.txt):

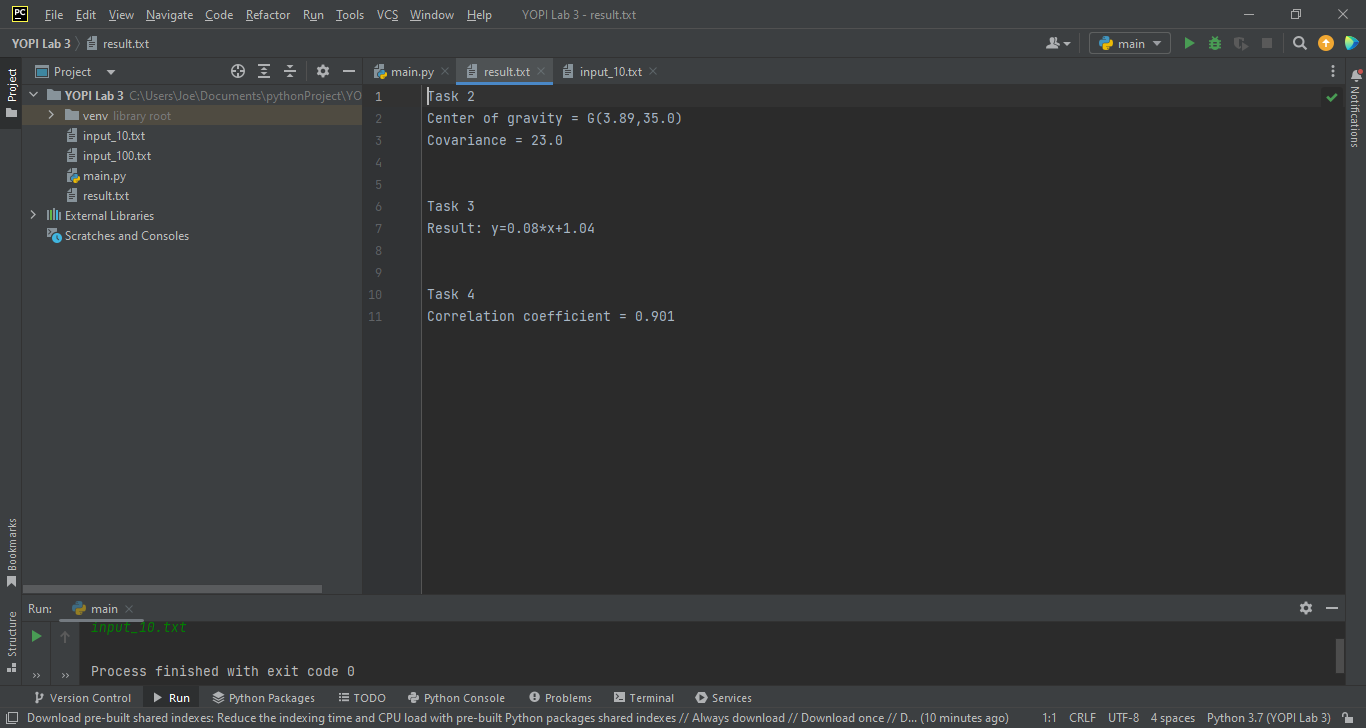


Задання 2:

Код:

sump=float(0)  
sumt=0  
cov=float(0)  
for i in range(0,paid.size):  
 sump+=paid[i]  
for i in range(0,time.size):  
 sumt+=time[i]  
middlep=float(sump/paid.size)  
middlet=float(sumt/time.size)  
file.write(f"Center of gravity = G({round(middlep,2)},{middlet}) \n")  
for i in range(0,time.size):  
 cov+=((paid[i]-middlep)\*(time[i]-middlet))/time.size  
file.write(f"Covariance = {round(cov,3)}\n")

Результат:

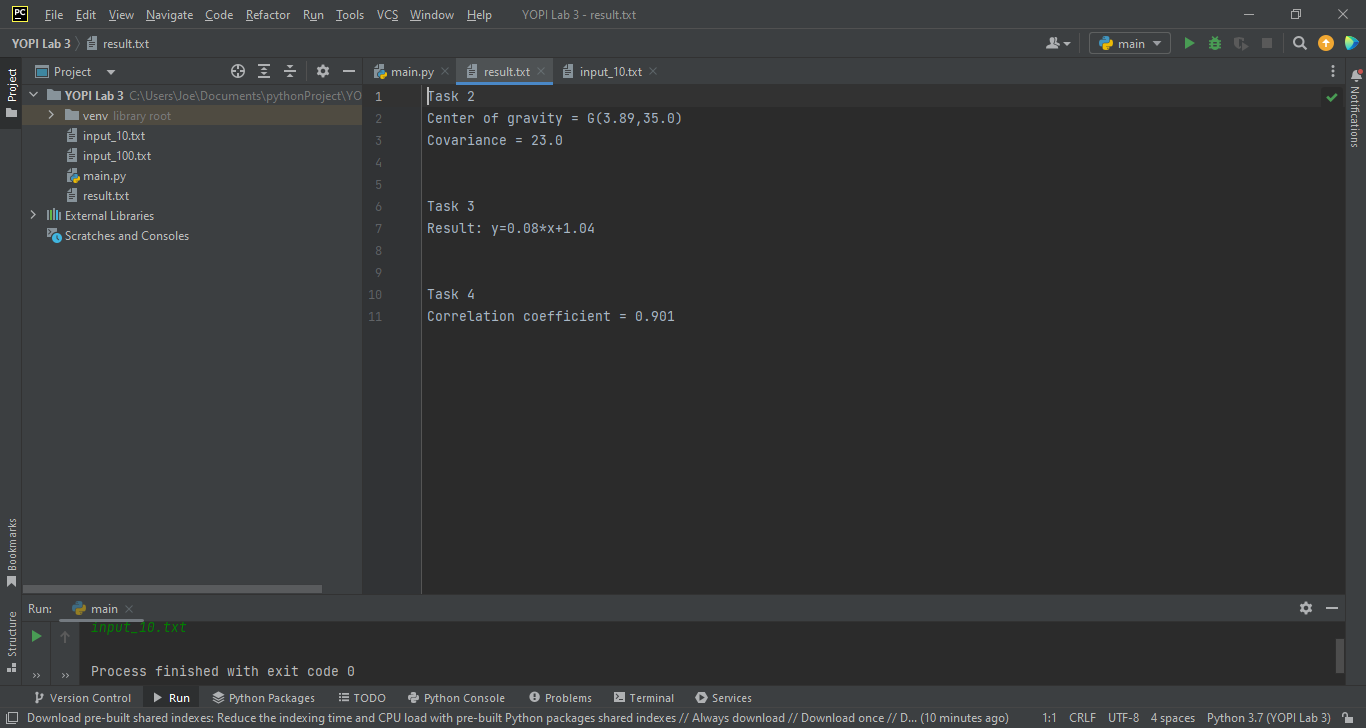


Завдання 3:

Код:

file.write("Task 3\n")  
var=float(0)  
for i in range(0,time.size):  
 var+=(math.pow(time[i],2)-math.pow(middlet,2))/time.size  
b1=float(cov/var)  
b0=float(middlep-b1\*middlet)  
file.write(f"Result: y={round(b1,2)}\*x+{round(b0,2)}\n")

Результат:

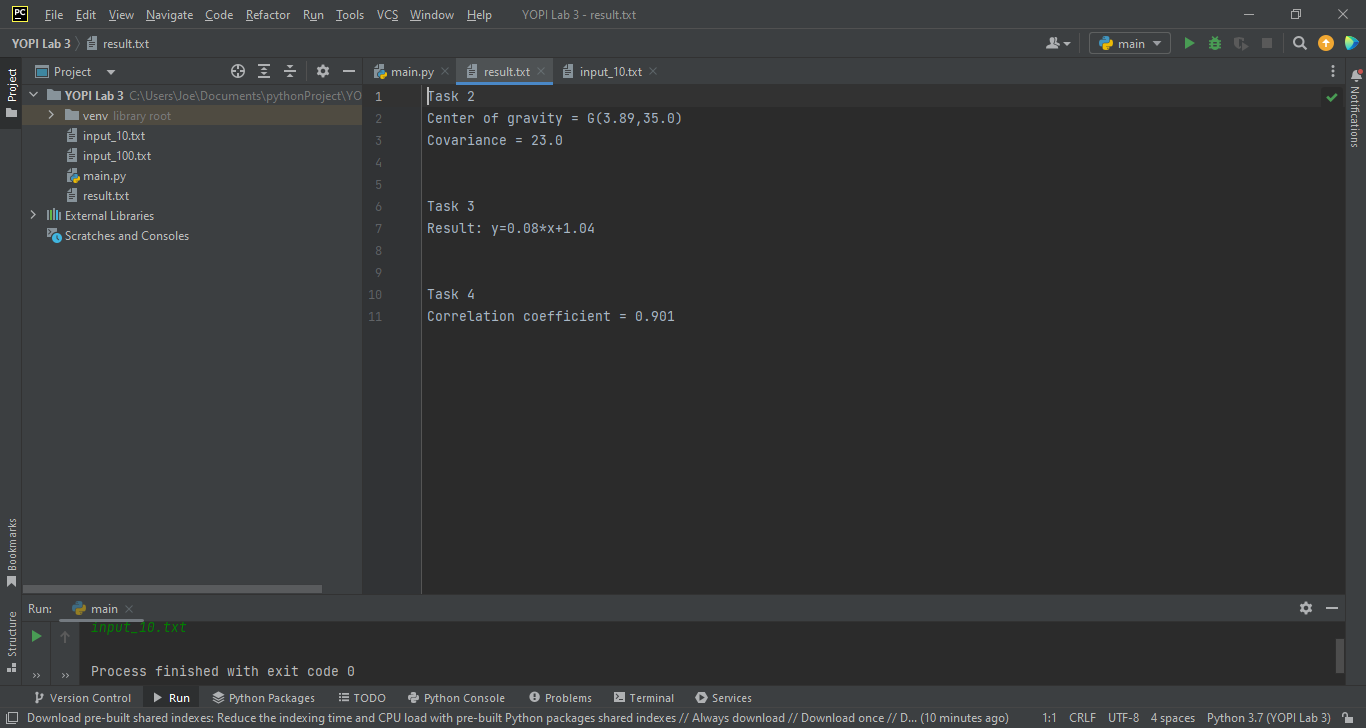


Завдання 4:

Код:

sumdift=float(0)  
sumdifp=float(0)  
devt=float(0)  
devp=float(0)  
cor=float(0)  
file.write("Task 4\n")  
for i in range(0,(time.size)):  
 sumdift+=math.pow((time[i]-middlet),2)  
devt=math.sqrt(sumdift/(time.size-1))  
for i in range(0,(paid.size)):  
 sumdifp+=math.pow((paid[i]-middlep),2)  
devp=math.sqrt(sumdifp/(paid.size-1))  
  
for i in range(0,(time.size)):  
 cor+=(((time[i]-middlet)/devt)\*((paid[i]-middlep)/devp))/(time.size-1)  
file.write(f"Correlation coefficient = {round(cor,3)}")

Результат:



**Висновки**

Під час виконання даної лабораторної роботи було використано набуті знання про двовимірну статистику. Було обраховано центр ваги, кореляцію, коефіціент кореляції та знайдено рівняння лінії регресії у від х.

На діаграмі розкиду даних бачимо, що сума покупки збільшується відповідно до часу, тож існує позитивна залежність. Також можна сказати, що дана залежність досить близька до лінійної.