Міністерство освіти і науки України Національний університет «Львівська політехніка»

Інститут телекомунікацій, радіоелектроніки та електронної техніки кафедра «Телекомунікацій»



Звіт з лабораторної роботи №1

з дисципліни «Теорія алгоритмів та структур даних»

Підготував:  
ст.групи ТР-31

Гейниш Р.Т.

Прийняв:

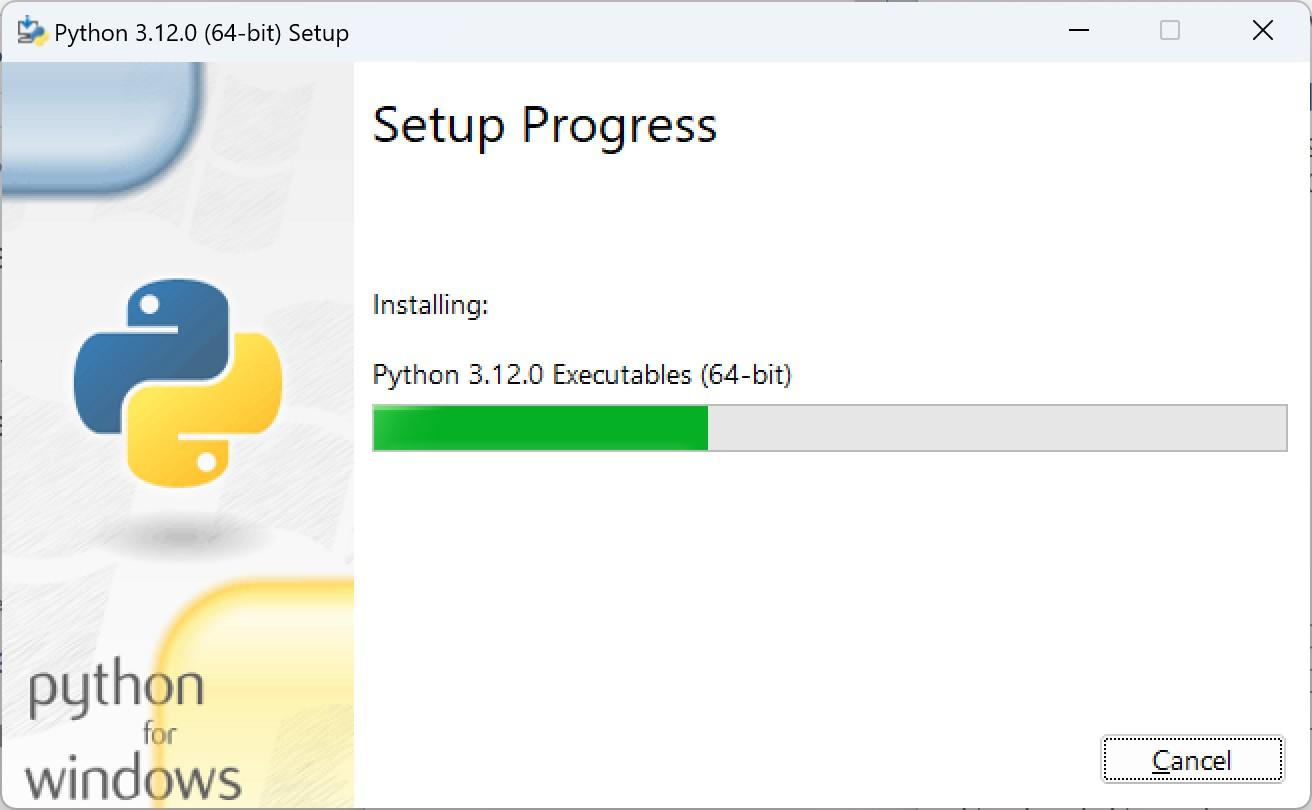
Андрущак В.С.

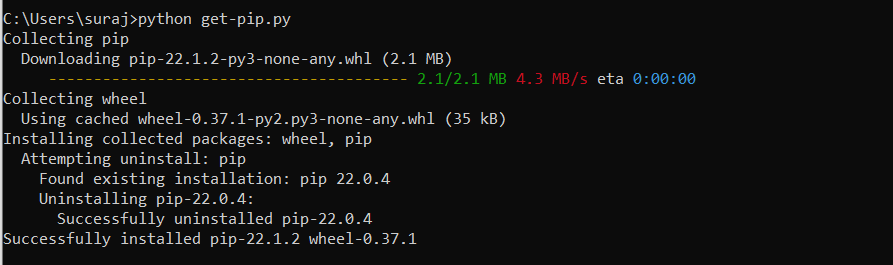
Львів 2024р.

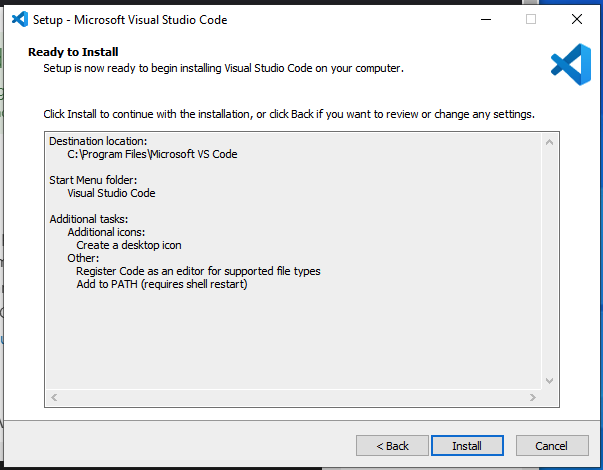
**Тема роботи:** Налаштування середовища роботи Python3 та Jupyter Notebook

**Мета роботи:** налаштувати робоче середовище, вивчити та дослідити основні технічні елементи для дослідження даних та алгоритм

# Завдання до виконання

**Завдання 1.** Налаштувати середовище роботи Встановлюємо Python3, пакетний менеджер pip (pip3) та VSCode.



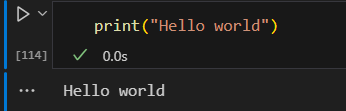


Піссля чого встановлюємо jupyter notebook, знайшовши його у вкладці extensions



За допомогою пакетного менеджера pip інсталяціюмо бібліотеки numpy, pandas, matplotlib ввівши pip install <назва бібліотеки>

Перевіряємо правильність налаштування середовища та бібліотек Підключаємо бібліотеки та виводимо у консоль “Hello, world!”

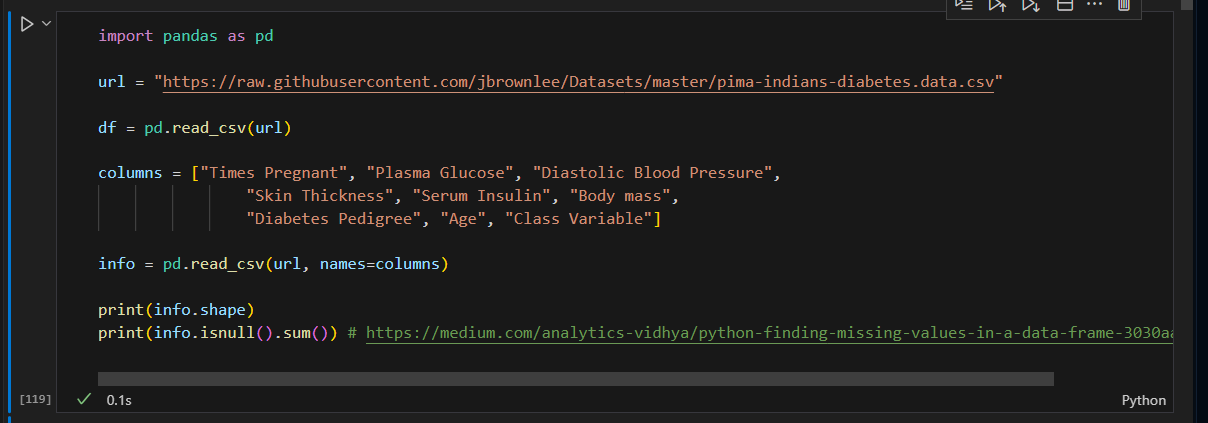


Генеруємо масив з 100 випадкових елементів та візуалізуємо дані

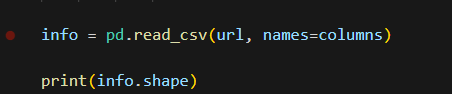


**Завдання 2.** Дослідження датасету

Підключаємо необхідні бібліотеки та завантажуємо датасет. Прочитавши опис до датасету називаємо відповідно колонки.



Визначаємо розмір датасету



Отримуємо наступне значення у виводі, де 768 - це кількість рядків, а 9 -

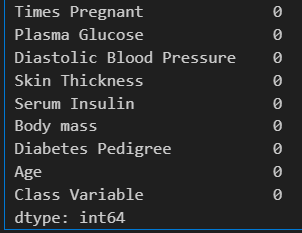
кількість стовпців:

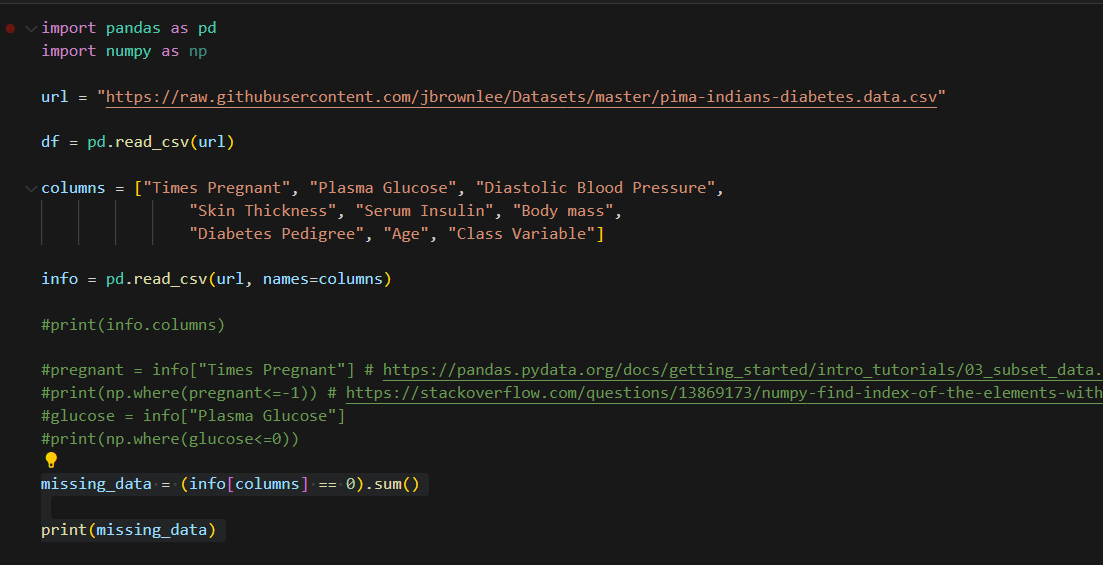
(768, 9))

Наступним кроком визначаємо пропущені дані. Для цього проведемо перевірку на пусті значення (NaN) за допо isnull() та виводимо загальну суму NaN значень з кожної колонки

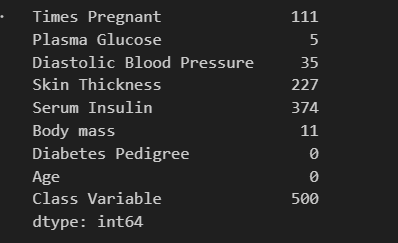


Отримуємо:



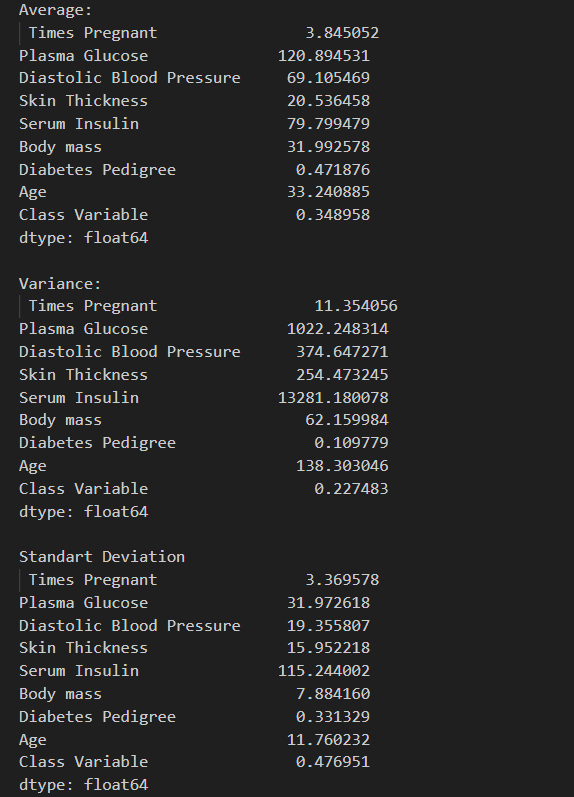
Отже, даних зі значенням NaN немає. А описі до датасету вказано, що пропущені дані є, перевіряємо колонки, які можуть містити значення “0”, але не є логічним, як наприклад колонка Plasma Glucose, що містить дані про концентрацію глюкози в плазмі через 2 години в оральному тесті на толерантність до глюкози, норма становить зазвичай 140 мг/д  


Отримуємо:

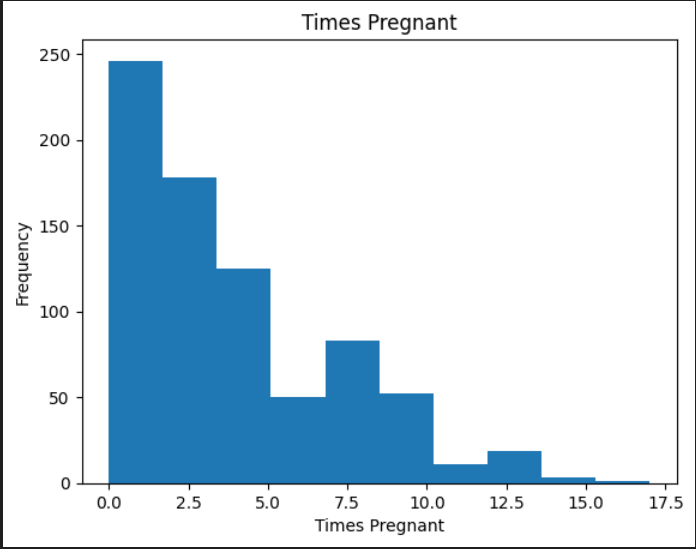


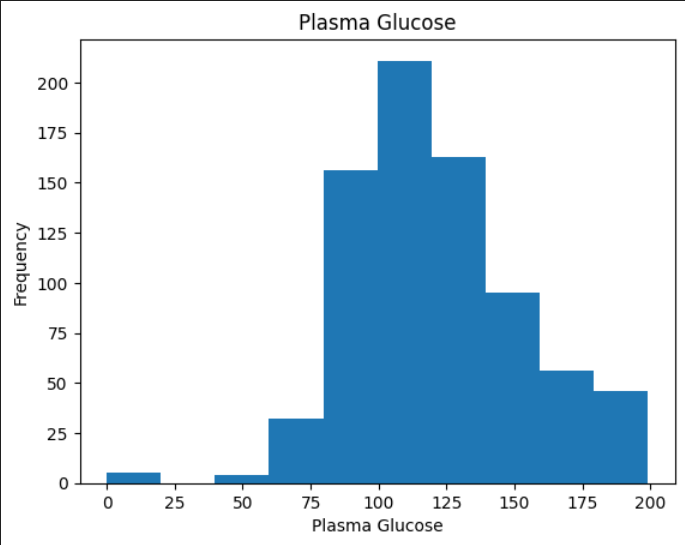
Здійснюємо обрахунок середнього арифметичного, дисперсії, середнє квадратичне відхилення для кожної із колонок набору даних:

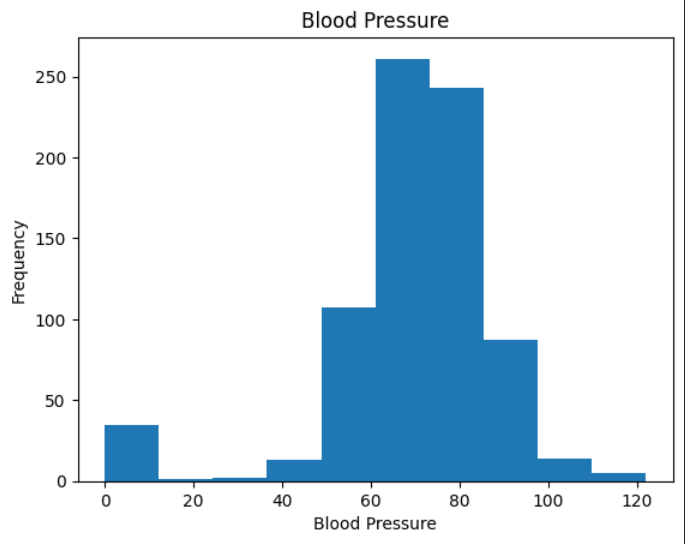


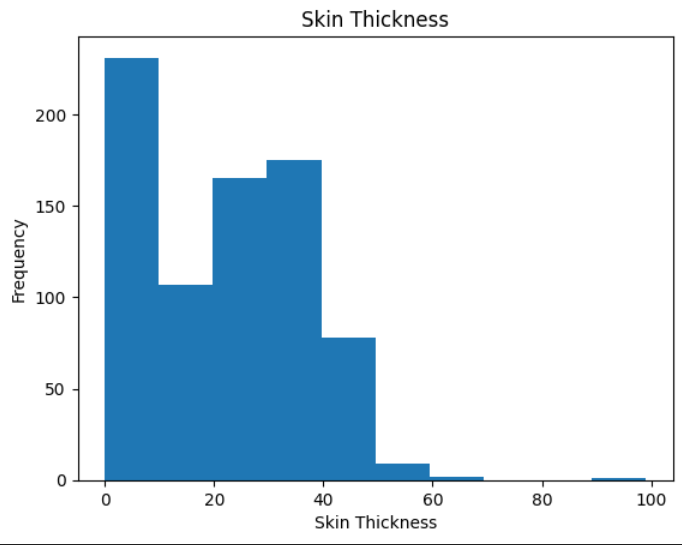
Отримуємо:  


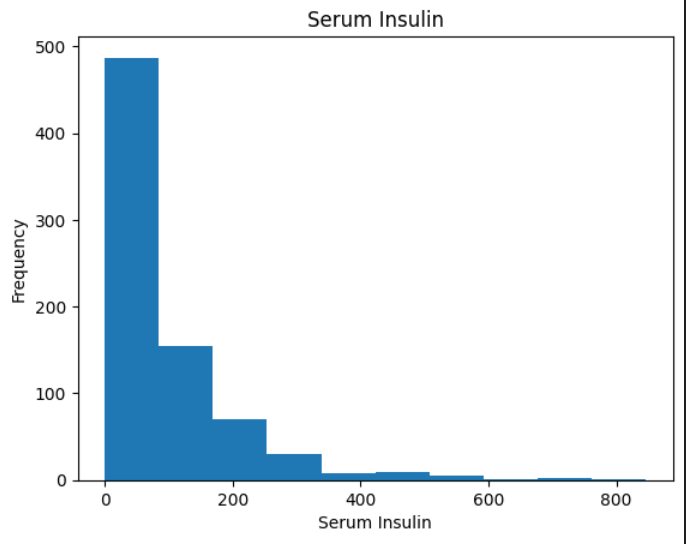
Візуалізємо дані за допомогою гістограм:

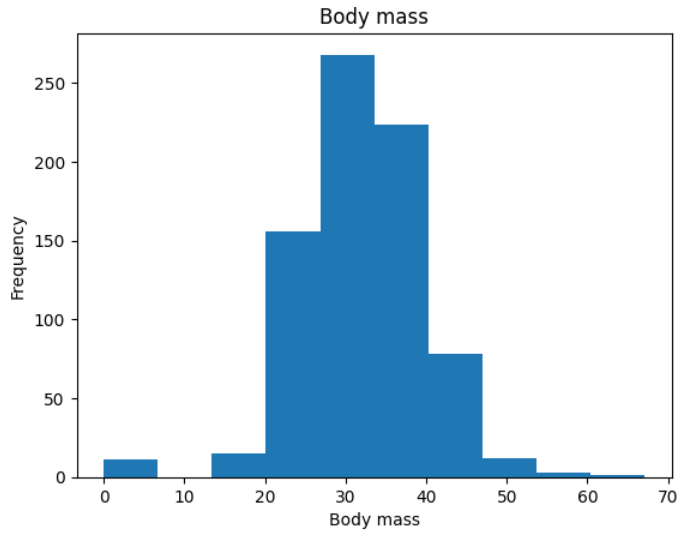


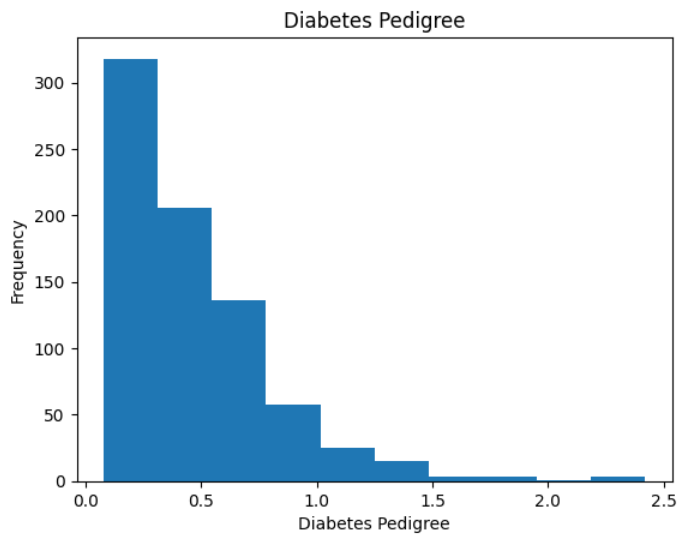


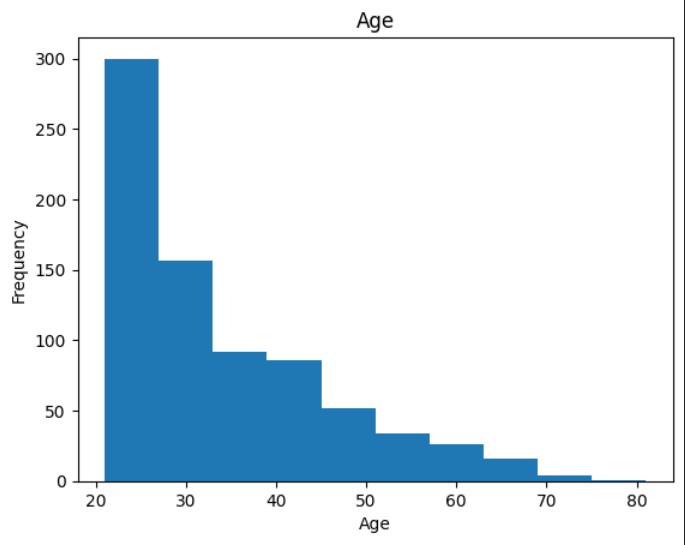


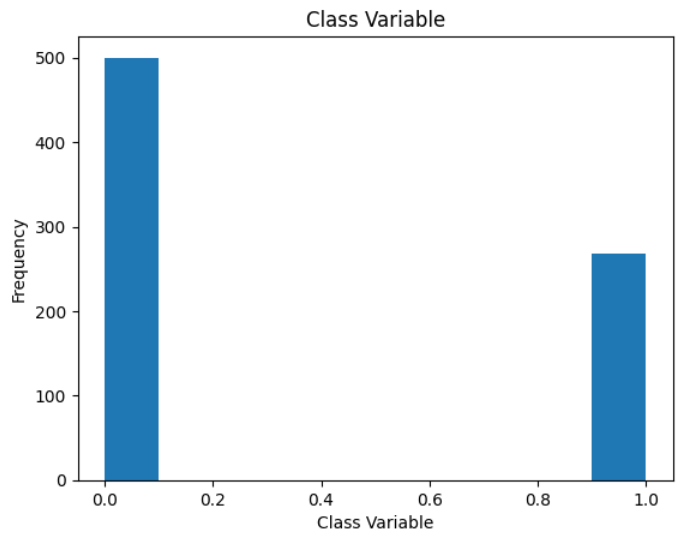












# Висновки:

Отже, на основі даних можна зробити загальні висновки. Можна помітити, що більшість пацієнтів мають менше ніж 5 вагітностей. Більшість пацієнтів також мають нормальне значення рівня глюкози крові, дані зосереджені в діапазоні від 100 до 150, (середнє значення становить 121.6) Дані про артеріальний тиск показують, що більшість пацієнтів мали тиск від 60 до 80. Розподіл товщини шкіри трицепса може бути корисним показником жирової маси, у більшості цей показник був рівний 20-40 мм.

Середнє значення індексу маси тіла (32.46) перевищує діапазон норми (18.5 до 24.9). Вибірка складалася з людей середній вік яких становить 33.2 роки. 500 людей мали негативний тест на діабет, 268 мали позитивний тест.