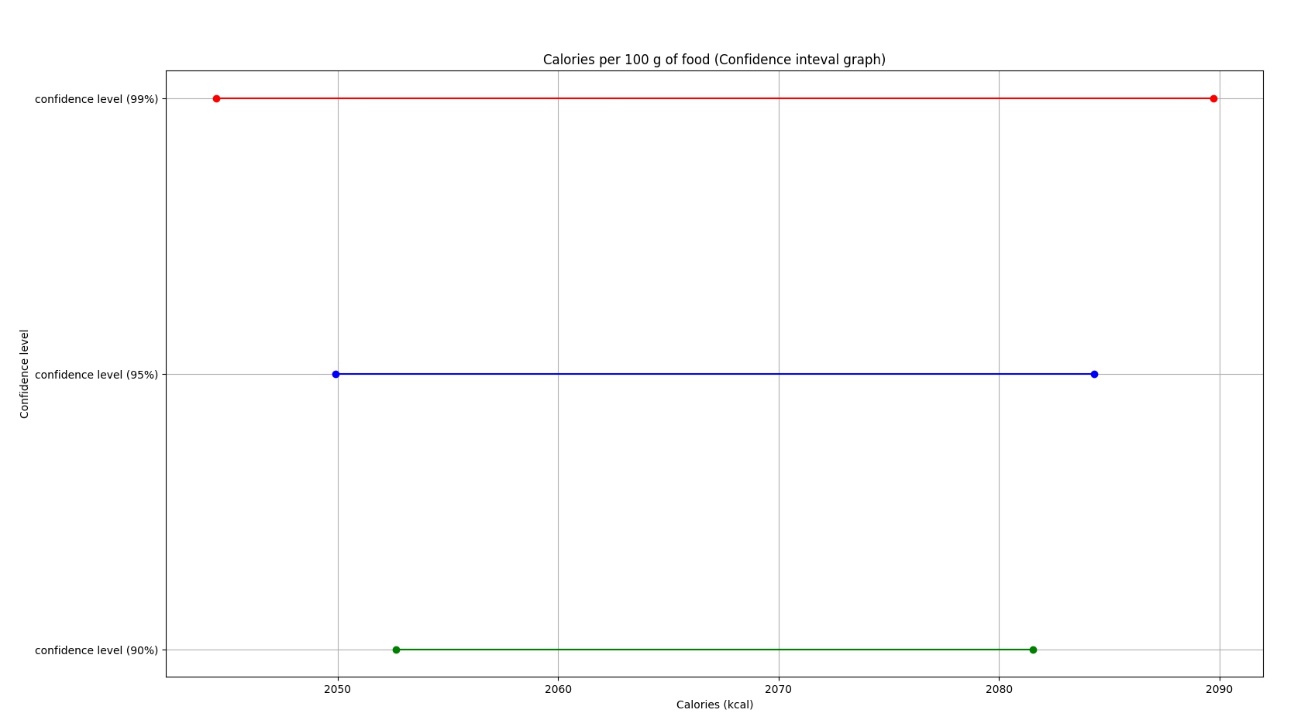
**นาย ภัทรพัทธิ์ ชัยอมรเวทย์ รหัสนักศึกษา : 62010684  
คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมคอมพิวเตอร์**

**Topic : Confidence Interval Graph (CI)**

**ภาษาที่ใช้ : python**

**Column ที่เลือกใช้ : Calories หน่วย (kcal)**

****

**ในแกน x ของกราฟ -> จะบอกถึงจำนวนของ Calories ในช่วงระดับความมั่นใจแต่ละช่วง ซึ่งมีหน่วยเป็น kcal**

**ในแกน y ของกราฟ -> จะบอกถึงระดับความมั่นใจที่มาใช้ในการคำนวณ ซึ่งแบ่งเป็น 3 ระดับ ได้แก่  
 90% , 95% และ 99%  
[ที่มาของกราฟ] -> ก่อนที่เราจะได้ Confidence interval graph มานั้น เราต้องรู้ก่อนว่าการทดลองของเรามี  
 องค์ประกอบอะไรบ้าง เช่น จำนวนประชากร, ค่าเออเร่อ, ค่าเฉลี่ย ฯลฯ ซึ่งระดับความเชื่อมั่น  
 จะคำนวณได้จากสูตรข้างล่างดังนี้  
 รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ**

**รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ**

**ในการคำนวณและการหาค่าข้อมูลเชิงสถิตินั้น เราสามารถเขียนโปรแกรมภาษา Python เพื่อที่จะนำมาคำนวณหาช่วงระดับความมั่นใจ (Confidence interval) ได้ โดยใช้ lib ของภาษานี้ ซึ่ง lib ที่เรานำมาใช้ได้แก่  
 1.matplotlib -> ใช้เพื่อวาดกราฟจากไฟล์ข้อมูล และจำแนกได้กราฟได้หลายรูปแบบ  
 2.pandas -> ใช้เพื่ออ่านข้อมูลที่เป็นตัวเลข (numeric data) จากไฟล์ csv ได้  
 3.statistic -> ใช้เพื่อคำนวณหาค่าข้อมูลเชิงสถิติต่างๆ เช่น ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น  
 4.math -> ใช้เพื่อคำนวณแล้วนำมาใช้กับสูตรระดับความมั่นใจ เช่น หารากที่สอง (sqrt)**

**Code : python**

**รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ**

**บรรทัดที่ 1-4 -> จะเป็นการ import lib เพื่อนำมาใช้งานด้านต่างๆ  
บรรทัดที่ 5-11 -> จะเป็นการดึงข้อมูลตัวเลขมาจากไฟล์ csv ซึ่งข้อมูลจะเป็นประเภทตัวเลข   
บรรทัดที่ 13-28 -> จะเป็นการคำนวณเชิงสถิติ โดยเราใช้ statistics lib เพื่อนำมาช่วยคำนวณ CI ด้วย  
บรรทัด 29-33 -> จะเป็นการกำหนดขอบล่างและขอบบนให้กับ CI แต่ละช่วงว่าถ้าระดับความมั่นใจช่วงนี้ควรมีจะมี  
 ขอบเขตเป็นเท่าไหร่ ใช้ค่า z เป็นเท่าไหร่ \*ค่า z สามารถเปิดตารางในเว็บดูได้\* โดย  
 CI 90% จะใช้ค่า z = 1.645 / CI 95% จะใช้ค่า z = 1.96 / CI 99% จะใช้ค่า z = 2.576  
บรรทัดที่ 35-44 -> จะเป็นการพลอตกราฟจากข้อมูลที่เราคำนวณมาในบรรทัดก่อนหน้า ซึ่งจะใช้ matplotlib   
 มาช่วยในการสร้างกราฟ**

**ผลลัพธ์ของโค้ด**

**รูปภาพประกอบด้วย ข้อความ

คำอธิบายที่สร้างโดยอัตโนมัติ**

**บทวิเคราะห์  
 จากกราฟ เราคำนวณออกมาแล้วได้ว่า ขอบเขตล่างและขอบเขตบนของแต่ละช่วงแต่ระดับความมั่นใจ  
มีกี่แคลรอลี่ ซึ่งจะสรุปได้ว่าถ้าระดับความมั่นใจอยู่ที่ 90% และเราสุ่มการทดลองไป 500 ครั้ง จะมีการครอบคลุม  
พารามิเตอร์อยู่ที่ 450 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มอยู่ในช่วง 2052.66 – 2081.55 และไม่มีการครอบคลุม  
พารามิเตอร์อยู่ที่ 50 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มไม่อยู่ในช่วง 2052.66-2081.55 ถ้าระดับความมั่นใจ  
อยู่ที่ 95 % และเราสุ่มการทดลองไป 500 ครั้ง จะมีการคลอบคลุมพารามิเตอร์ จะมีการครอบคลุมพารามิเตอร์อยู่ที่  
475 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มอยู่ในช่วง 2049.89 – 2084.32 และไม่มการรอบคลุมพารามิเตอร์อยู่ที่  
25 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มไม่อยู่ในช่วง 2049.89 – 2084.32 ถ้าระดับความมั่นใจอยู่ที่ 99% และเรา  
สุ่มการทดลองไป 500 ครั้ง จะมีการครอบคลุมพารามิเตอร์อยู่ที่ 495 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มอยู่  
ในช่วง 2044.48 – 2089.73 และไม่มีการครอบคลุมพารามิเตอร์อยู่ที่ 5 ครั้ง ก็คือ จำนวนแคลรอลี่จากการสุ่มไม่อยู่ในช่วง 2044.48 – 2089.73 ซึ่งค่าที่เราสร้างขึ้นนั้นจะเป็นการทดลองที่เป็นแบบช่วงนั่นเอง**