

1. 請將資料 hwddata.csv 匯入 R，並檢視資料的內容，請問這筆資料收集了幾個人、幾個變數
2. 請計算每個人的 BMI，並將其四捨五入到小數點後第 3 位
3. 請建立一個資料，只保留性別 (sex)、年齡 (age)、BMI、脈搏 (pulse) 等 4 個變數
4. 請使用 3.建立的資料，計算男女生分別有多少人，男女分別的平均 BMI 為多少
5. 請使用 3.建立的資料，利用脈搏 (pulse) 來建立新變數 heart，用以代表心跳是否過快，若  $\text{pulse} \geq 90$  則  $\text{heart} = 1$ 、 $\text{pulse} < 90$  則  $\text{heart} = 0$ 。篩選資料只留下心跳過快的人，請問心跳過快共幾人？
6. 請以所有人的資料，計算體重(Weight)及三酸甘油脂(TG)之描述性統計，使用下表描述此兩個變項的分布情況

---

樣本數(N) 平均數(mean) 標準差(sd) 中位數(median) 四分位差(IQR)
體重(Weight)
三酸甘油脂(TG)

---

7. 請以性別(Sex)分組，分別計算男生的體重(Weight)及三酸甘油脂(TG)之描述性統計，使用下表描述性別此兩個變項的分布情況。

---

樣本數(N) 平均數(mean) 標準差(sd) 中位數(median) 四分位差(IQR)
體重(Weight)
三酸甘油脂(TG)

---

8. 請先計算 BMI 並將其分為三組：第 1 組體重過輕 ( $\text{BMI} < 18.5$ )、第 2 組健康體位 ( $18.5 \leq \text{BMI} < 24$ )，及第 3 組過重和肥胖 ( $24 \leq \text{BMI}$ )。請計算不同性別中，各組 BMI 的人數以及所占百分比。請輸出所需的百分比就好，若輸出總百分比不算分。

9. 承上題，若想呈現不同性別「BMI」的分佈需要使用哪一種統計圖形？為什麼？最後請繪製出該統計圖形
10. 承上題，若想呈現不同性別「BMI 分組」的分佈需要使用哪一種統計圖形？為什麼？最後請繪製出該統計圖形
11. 假設此資料為一個台灣成年人的隨機樣本，且具代表性，現欲檢定台灣成年人的心跳(Pulse) 平均值是否為 80 次/分鐘 (給定顯著水準  $\alpha=0.05$ ，請執行雙尾檢定)？
12. 承上題，請檢定男、女性的心跳(Pulse)平均值是否相等 (給定顯著水準  $\alpha=0.05$ ，請執行雙尾檢定)？
13. 承上題，請先將年齡(Age)分為三組：第一組( $\text{Age} \leq 30$ )、第二組( $30 < \text{Age} \leq 50$ )、第三組( $50 < \text{Age}$ )，並以 ANOVA 檢定不同年齡分組膽固醇(Cholesterol)平均值是否相等，如果有不同，請問是哪幾組不同？最後，請說明此情況下使用 ANOVA 檢定是否合理，原因為何？