**流行病學 實習課 SAS 作業一**

公衛二

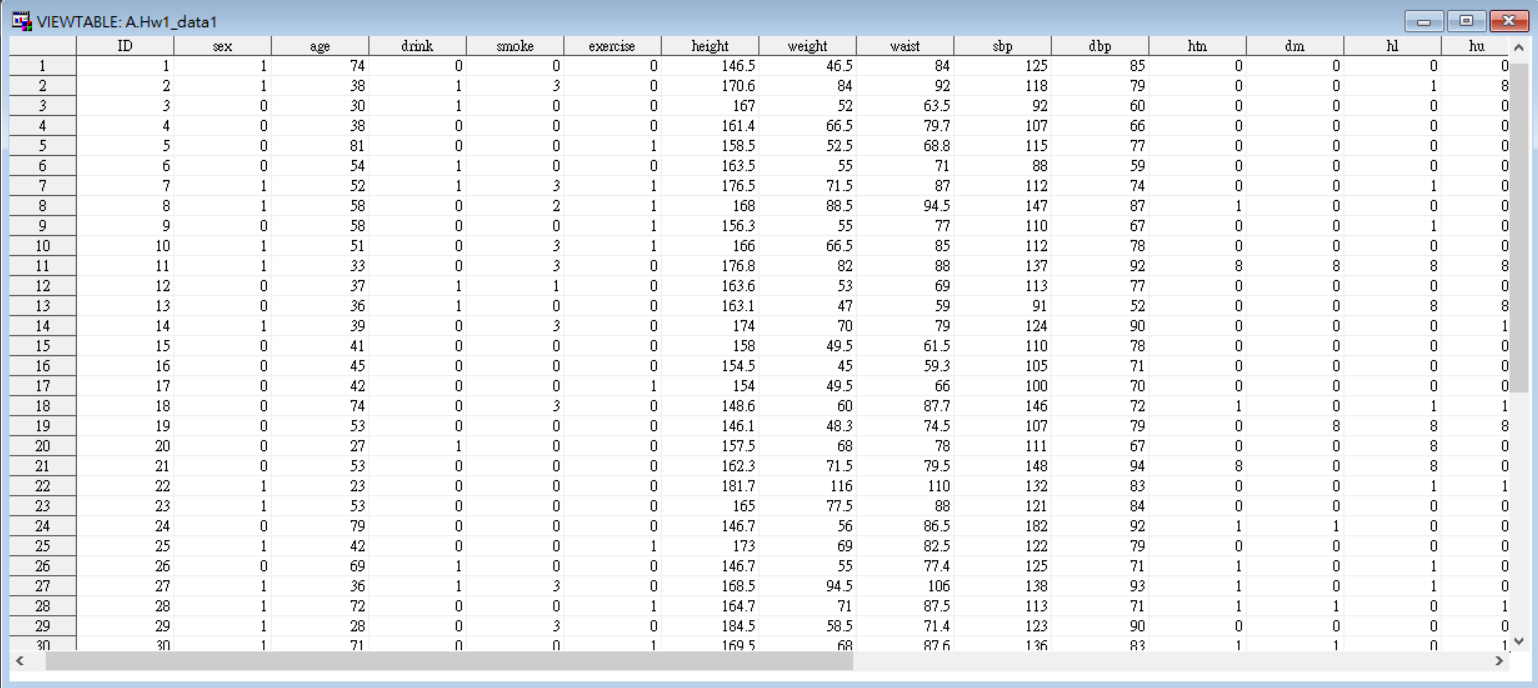
梁嫚芳

B07801003

繳交日期: 2020.04.14

1. **某社區隨機抽取 50 人的健康調查資料，該資料分別被存於 HW1\_basic 和HW1\_health 檔案中。請利用 SAS 程式語言將兩個檔案 HW1\_basic 和HW1\_health 合併後存於一個永久資料庫中。(8%)**

建立永久資料庫”A”，將HW1\_basic和hw1\_health匯入，分別命名為hw1\_basic與hw1\_health，將兩者水平合併並存於永久資料庫”A”中，命名為hw1\_data1(如圖)



/\*1\*/

/\*建立永久資料庫’a’\*/

libname a 'D:\SAS' ;

/\*SORT\*/

PROC SORT data=a.Hw1\_basic;

by ID ;

RUN;

PROC SORT data=a.Hw1\_health;

by ID ;

RUN;

/\*合併兩組資料\*/

**DATA** a.hw1\_data1;

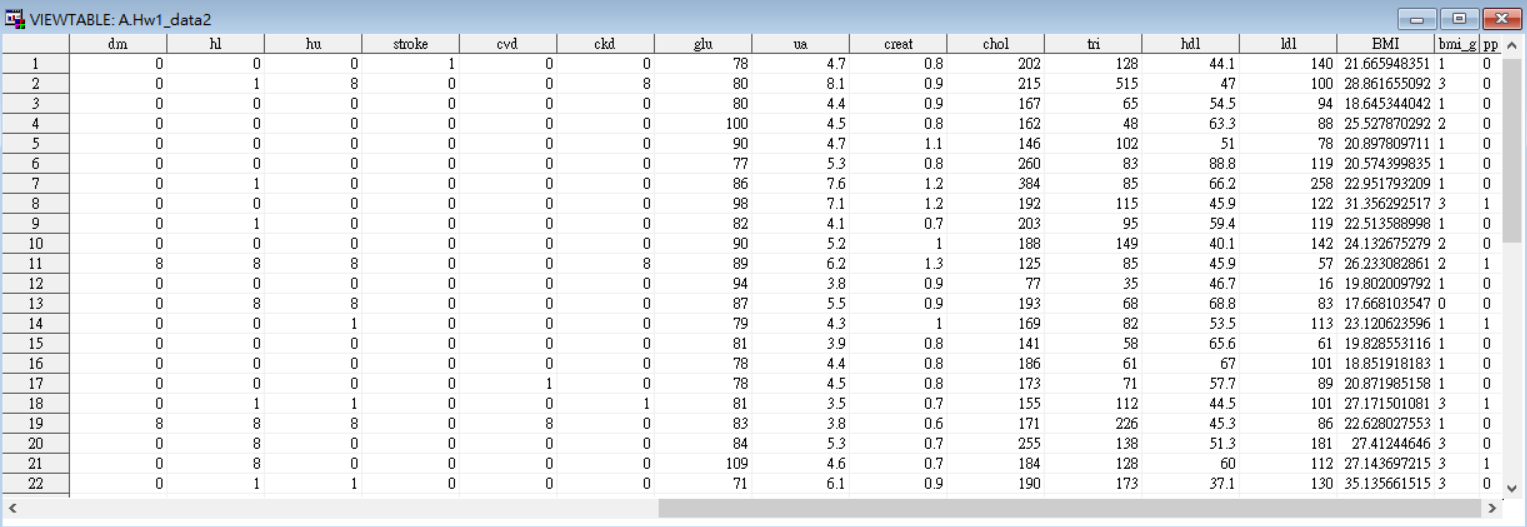
merge a.Hw1\_basic a.Hw1\_health;

by ID;

**RUN**;

1. **為方便檢視某社區居民之健康狀況，將每位居民的 BMI、血壓依照衛服部之標準分組檢視。BMI 分為四組：過輕(<18.5)、正常(18.5≦BMI<24)、過重(24≦BMI<27)、肥胖(≧27)。血壓分為正常(SBP<140mmHg 和DBP<90mmHg)與異常兩組。 (20%)**

建立新資料檔”Hw1\_data2”，計算bmi(體重除以身高平方)後建立新變項”BMI”，將BMI分為四組(過輕:0、正常:1、過重:2、肥胖:3)建立為新變項”bmi\_g”，將血壓分為兩組(正常:0、異常:1)建立為新變項”pp”。(如圖)



/\*2\*/

**DATA** a.hw1\_data2; set a.hw1\_data1;

/\*bmi分組\*/

BMI=weight/((height/**100**)\*\***2**);

if BMI<**18.5** then bmi\_g='0';

else if BMI<**24** then bmi\_g='1';

else if BMI<**27** then bmi\_g='2';

else bmi\_g='3';

/\*血壓分組\*/

if sbp<**140** and dbp<**90** then pp='0';

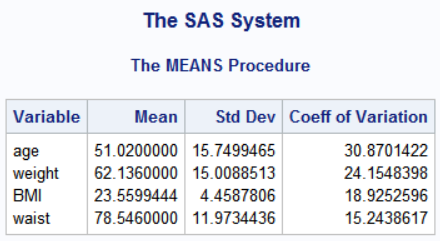
else pp='1';

**RUN**;

1. **推行預防醫學，社區健康中心之公衛護理師需透過該社區居民的年齡、體重、BMI 和腰圍來評估居民的肥胖狀況，請協助計算上述變項的平均值、標準差及變異係數。(12%)**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 平均值 | 標準差 | 變異係數 |
| 年齡 | 51.02 | 15.7499465 | 30.8701422 |
| 體重 | 62.136 | 15.0088513 | 24.1548398 |
| BMI | 23.5599444 | 4.4587806 | 18.9252596 |
| 腰圍 | 78.546 | 11.9734436 | 15.2438617 |

Output:



/\*3\*/

/\*計算平均值 標準差 變異數 \*/

**PROC** **MEANS** DATA=a.Hw1\_data2 mean std cv;

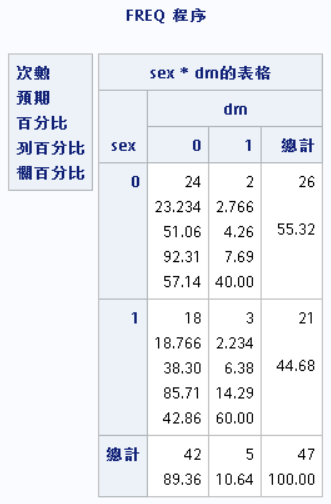
VAR age weight BMI waist ;

**RUN**;

1. **接著，預計推動健康促進介入計畫，公衛護理師調查該社區男性與女性的糖尿病病史分布情形，我們想知道性別與患有糖尿病是否有相關。請檢定虛無假設並說明你的結論為何？ (20%)**

H0: 性別與是否患有糖尿病無相關

H1: 性別與是否患有糖尿病有相關

 根據 chi-square test ，50%細胞個數期望值<5，因此改以Fisher’s exact test，p-value=0.6441> α(0.05)，不拒絕H0，因此性別與是否患有糖尿病沒有統計顯著的相關性。

/\*4\*/

/\*刪除不知道是否患有糖尿病的資料\*/

**DATA** a.Hw1\_data4; set a.Hw1\_data2;

if dm=**8** then delete;

**run**;

/\*卡方檢定\*/

**PROC** **FREQ** DATA=a.Hw1\_data4;

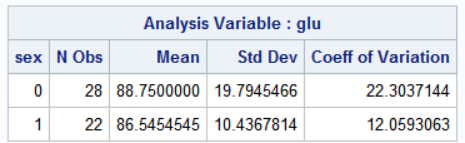
TABLE sex\*dm /expected chisq ;

**RUN**;

1. **根據上述的調查結果，社區健康中心想進一步確認男性和女性血糖之平均值、標準差、變異係數。請計算上述數值，並確認平均血糖值在不同性別間是否一樣呢？ (30%)**

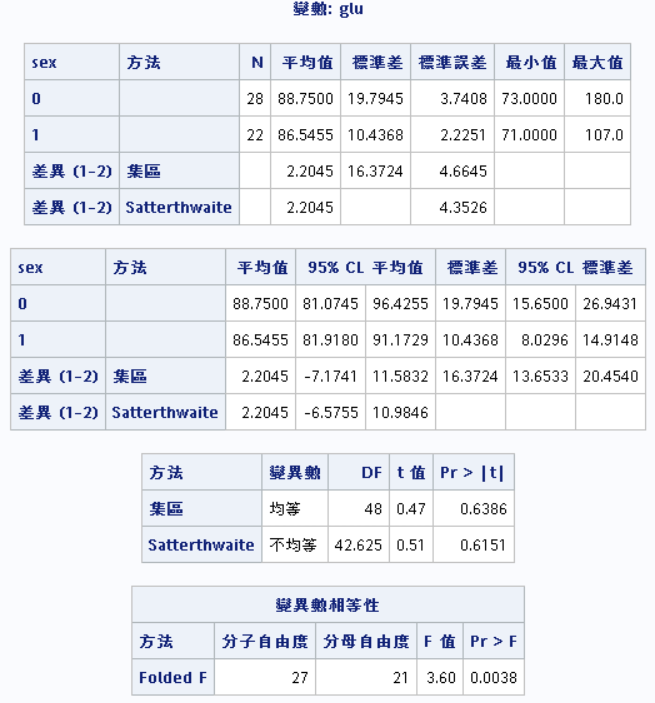
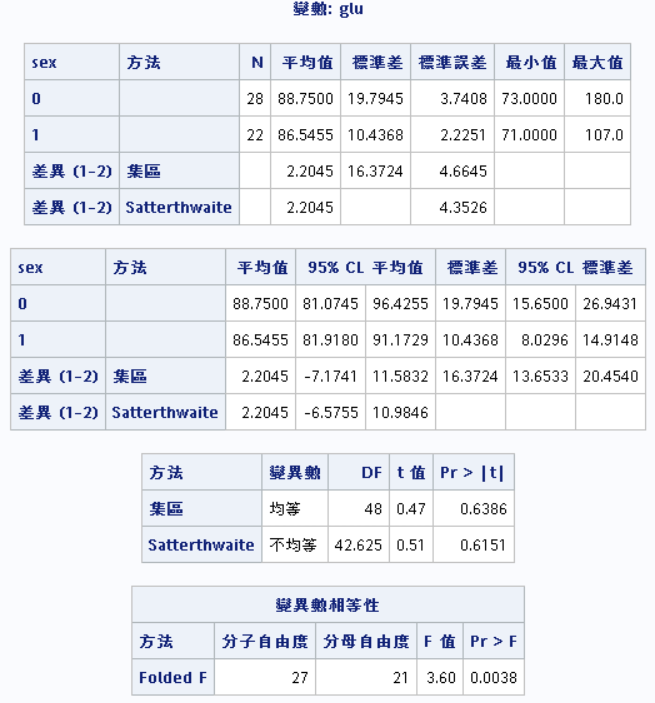
Output:

|  |  |  |  |
| --- | --- | --- | --- |
|  | 平均值 | 標準差 | 變異係數 |
| 女血糖 | |  | | --- | | 88.7500000 | | |  | | --- | | 19.7945466 | | |  | | --- | | 22.3037144 | |
| 男血糖 | |  | | --- | | 86.5454545 | | |  | | --- | | 10.4367814 | | |  | | --- | | 12.0593063 | |



平均血糖值就上表數值而言在不同性別間不相同。

但當進行兩組獨立樣本t-test，H0:平均血糖值在不同性別間相等，根據以下報表p-value(變異數不均等→Welch-Satterthwaite t-test)=0.6151> α(0.05)，因此不拒絕H0，平均血糖值在不同性別間無統計顯著的差異。



/\*5\*/

/\*以性別分組計算平均值 標準差 變異數\*/

**PROC** **MEANS** DATA=a.Hw1\_data2 mean std cv;

CLASS sex;

VAR glu ;

**RUN**;

/\*兩組獨立樣本t-test\*/

**PROC** **TTEST** DATA=a.Hw1\_data2;

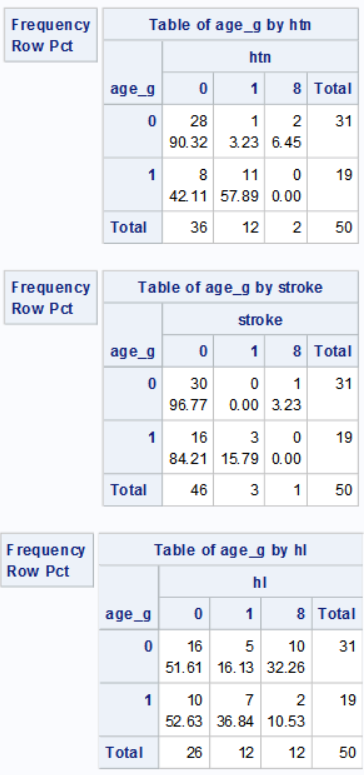
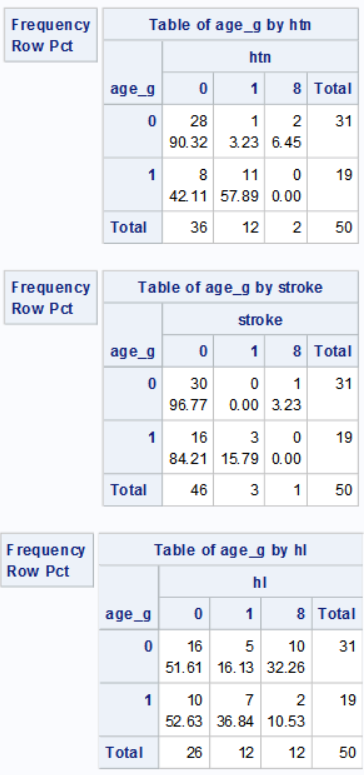
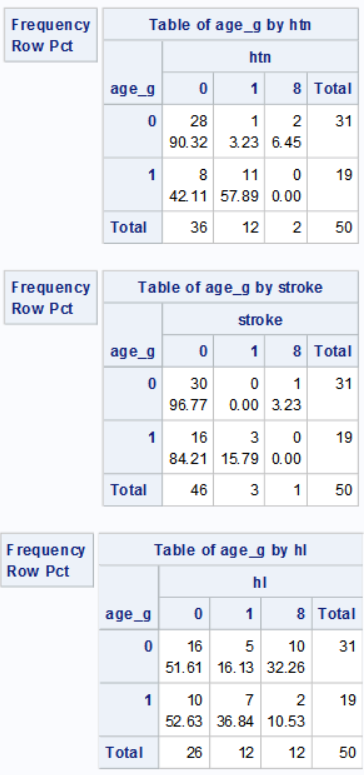
CLASS sex;

VAR glu;

**RUN**;

1. **社區健康中心乃重視除糖尿病以外的其他慢性疾病，請協助計算高血壓、中風、高血脂在年齡小於 55 歲和大於等於 55 歲之盛行率為何？ 以一種合適的圖形表示，並列出盛行率。 (10%)**

Output:



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 年齡 | 高血壓 | 中風 | 高血脂 |
| 盛行率(%) | <55 | 3.23 | 0 | 16.13 |
| ≥55 | 57.89 | 15.79 | 36.84 |

|  |  |  |
| --- | --- | --- |
| 年齡 | <55 | ≥55 |
| 高血壓 |  |  |
| 中風 |  |  |
| 高血脂 |  |  |

/\*6\*/

/\*將年齡分組 \*/

**DATA** a.Hw1\_data6; set a.Hw1\_data2;

if age<**55** then age\_g=**0**;

else age\_g=**1**;

**RUN**;

/\*計算盛行率 \*/

**PROC** **FREQ** DATA=a.Hw1\_data6;

TABLE age\_g\*htn age\_g\*stroke age\_g\*hl /nopercent nocol;

**RUN**;

/\*將資料以年齡組別排序 \*/

**PROC** **SORT** DATA=a.Hw1\_data6;

by age\_g;

**RUN**;

/\*圓餅圖\*/

**PROC** **GCHART** DATA=a.Hw1\_data6;

By age\_g;

PIE htn stroke hl/type = percent ;

**RUN**;

**QUIT**;