111 學年度第二學期科學計算軟體作業一

系級：114 姓名：黃薇庭 學號：F64101032

※ 書面作業格式為 pdf 檔，檔名：HW1_學號.pdf

※ 將程式碼附在作業最後

1. 請算出 2009-2011 三年間全台溫度之最大值、最小值、平均值、標準差以及四分位數(每一項 3 分共 15 分；請將各數值四捨五入取到小數第二位，**若未達到則會斟酌扣分**)。

   ANS:

```
> #Q1
> #max
> round(max(dataset$temperature,na.rm = TRUE),2)
[1] 31.09
> #min
> round(min(dataset$temperature,na.rm = TRUE),2)
[1] 7.43
> #mean
> round(mean(dataset$temperature,na.rm=TRUE),2)
[1] 23.64
> #Standard Deviation
> round(sd(dataset$temperature,na.rm = TRUE),2)
[1] 4.82
> #quantile
> round(quantile(dataset$temperature,0.25,na.rm = TRUE),2)
 25%
19.8
> round(quantile(dataset$temperature,0.50,na.rm = TRUE),2)
  50%
24.26
> round(quantile(dataset$temperature,0.75,na.rm = TRUE),2)
  75%
28.21
```

2. 請算出 2009-2011 每一年全台溫度之最大值、最小值、平均值、標準差以及四分位數(每一項 3 分共 45 分；請將各數值四捨五入取到小數第二位，**若未達到則會斟酌扣分**)。

   ANS:

```
> #Q2
> #max
> round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=max),2)
  type     x
1 2009 30.92
2 2010 31.09
3 2011 31.08
> #min
> round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=min),2)
  type     x
1 2009 10.04
2 2010 11.06
3 2011  7.43
> #mean
> round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=mean),2)
  type     x
1 2009 23.94
2 2010 23.69
3 2011 23.28
> #standard deviation
> round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=sd),2)
  type    x
1 2009 4.69
2 2010 4.49
3 2011 5.22
> #quantile
> round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=quantile,probs=c(0.25,0.50,0.7
5)),2)
  type x.25% x.50% x.75%
1 2009 20.33 24.33 28.55
2 2010 20.08 24.07 27.92
3 2011 18.13 24.30 28.30
```

3. 請畫出 2009-2011 三年間**臺南市地區不同季節**下溫度之直方圖(10 分；請依照 2 column, 2 row 呈現，X 軸命名為 Temperature(℃)，Y 軸命名為 Count；**答題提醒:12-2 月為冬季，3-5 月春季，6-8 月為夏季，9-11 月為秋季，不用根據季節取平均值來畫圖**；以上若未達到皆會斟酌扣分)。

   <mark>ANS:</mark>

   (有先在 excel 上將資料按照四季分類，並區分出台南市和其他縣市，存成 Tainan_season.csv)
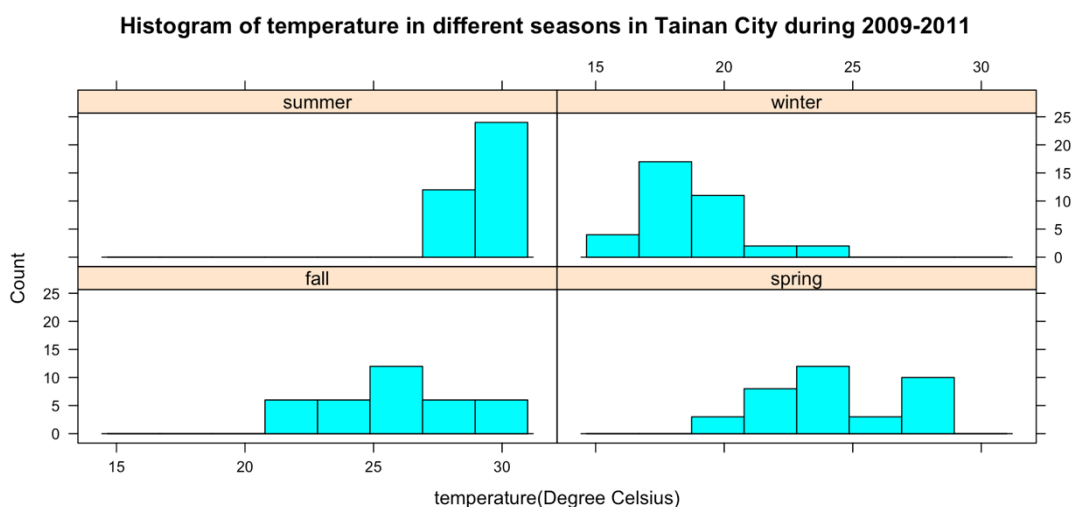
4. 請分別畫出 2009-2011 每一年**南市地區不同季節**下溫度之直方圖(30 分；請依照 2 column, 2 row 呈現，X 軸命名為 Temperature()，Y 軸命名為 Count；**答題提醒:12-2 月為冬季，3-5 月春季，6-8 月為夏季，9-11 月為秋季，不用根據季節取平均值來畫圖**；以上若未達到皆會斟酌扣分)。
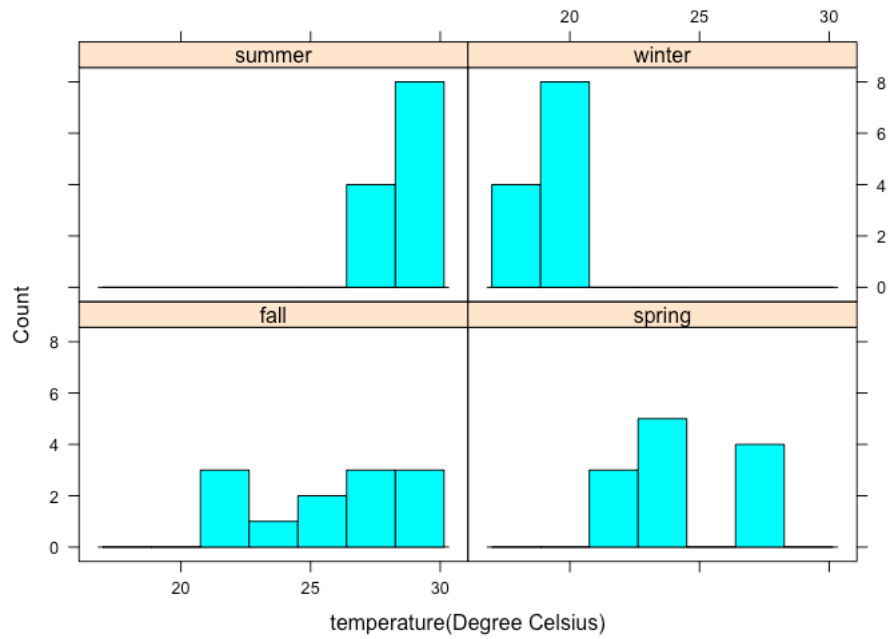
<mark>ANS:</mark>
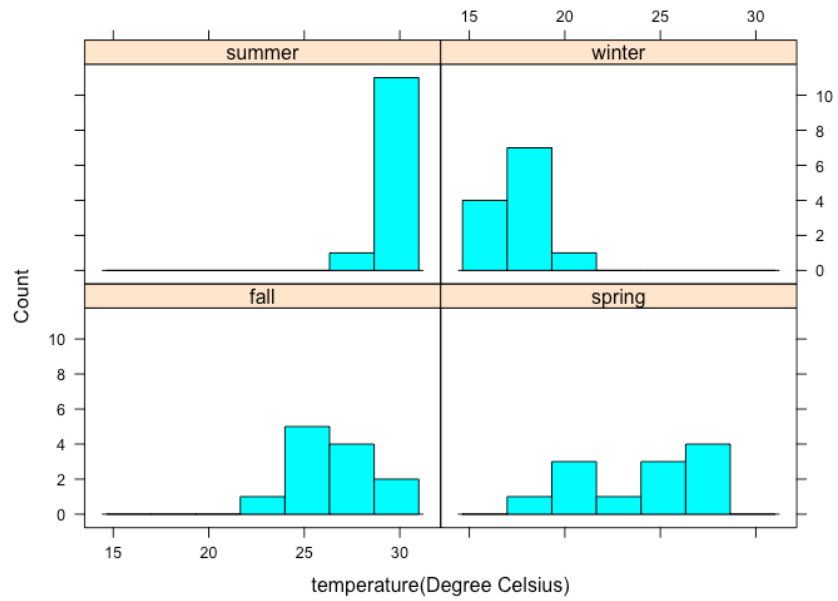
(有先在 excel 上將資料按照四季分類，並區分出台南市和其他縣市，按照年份分別存成

Tainan_season_2009.csv/Tainan_season_2010.csv/Tainan_season_2011.csv)

Histogram of temperature in different seasons in Tainan in 2009
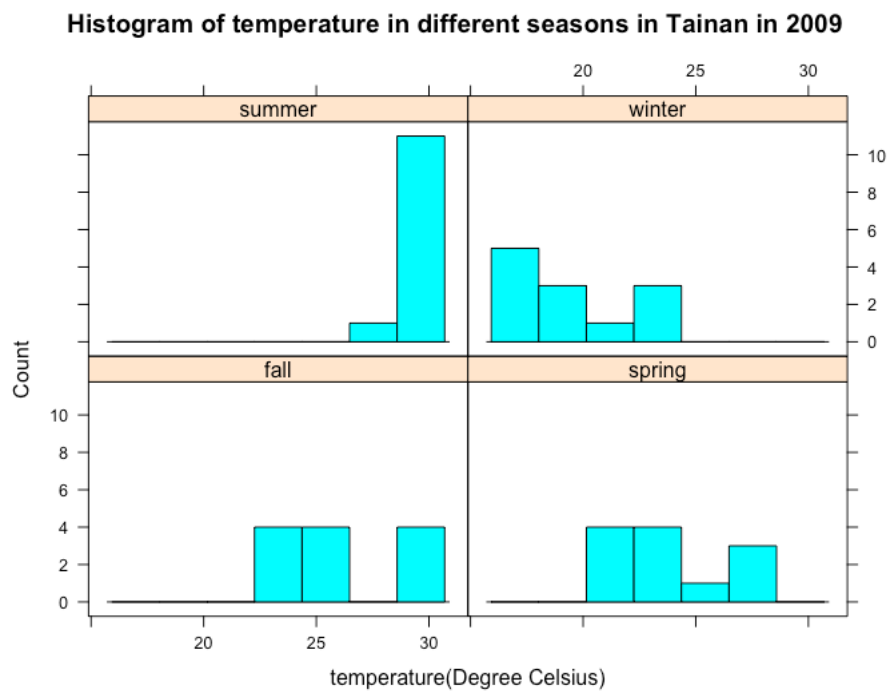
Histogram of temperature in different seasons in Tainan in 2010

Histogram of temperature in different seasons in Tainan in 2011

※　程式碼

```
###########################
# Set Working Directory #
###########################

# Get your current working directory #
getwd() # Re-check the path for the working directory

# Change your current working directory #
setwd("/Users/huangweiting/coding/INTRODUCTION TO SCIENTIFIC COMPUTING SOFTWARE
/C2_ClassData")
getwd() # Re-check the path for the working directory

#############################
# Import an example dataset #
#############################

dataset <- read.csv("C2_HW.csv")
#str(dataset)      #Check the variable format 看資料整體的data
#View(dataset)            #Check Dataset
#summary(dataset)         #Get the summary statistics about the object 給基礎統計數據（和資料
格式有關）

#Q1
#max
round(max(dataset$temperature,na.rm = TRUE),2)
#min
round(min(dataset$temperature,na.rm = TRUE),2)
#mean
round(mean(dataset$temperature,na.rm=TRUE),2)
#Standard Deviation
round(sd(dataset$temperature,na.rm = TRUE),2)
#quantile
round(quantile(dataset$temperature,0.25,na.rm = TRUE),2)
round(quantile(dataset$temperature,0.50,na.rm = TRUE),2)
round(quantile(dataset$temperature,0.75,na.rm = TRUE),2)

#Q2
#max
round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=max),2)
#min
round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=min),2)
#mean
round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=mean),2)
#standard deviation
round(aggregate(dataset$temperature, by=list(type=dataset$year),FUN=sd),2)
#quantile
round(aggregate(dataset$temperature,
by=list(type=dataset$year),FUN=quantile,probs=c(0.25,0.50,0.75)),2)

#Q3
dataset_2 <- read.csv("Tainan_season.csv")
histogram(x= ~temperature|season,data =dataset_2, xlab = "temperature(Degree
Celsius)",ylab = "Count",type="count",layout=c(2,2),main='Histogram of temperature in
different seasons in Tainan during 2009-2011')

#Q4
dataset_2 <- read.csv("Tainan_season_2009.csv")
histogram(x= ~temperature|season,data =dataset_2, xlab = "temperature(Degree
Celsius)",ylab = "Count",type="count",layout=c(2,2),main='Histogram of temperature in
different seasons in Tainan in 2009')
dataset_2 <- read.csv("Tainan_season_2010.csv")
histogram(x= ~temperature|season,data =dataset_2, xlab = "temperature(Degree
Celsius)",ylab = "Count",type="count",layout=c(2,2),main='Histogram of temperature in
different seasons in Tainan in 2010')
dataset_2 <- read.csv("Tainan_season_2011.csv")
histogram(x= ~temperature|season,data =dataset_2, xlab = "temperature(Degree




Celsius)",ylab = "Count",type="count",layout=c(2,2),main='Histogram of temperature in
different seasons in Tainan in 2011')
```