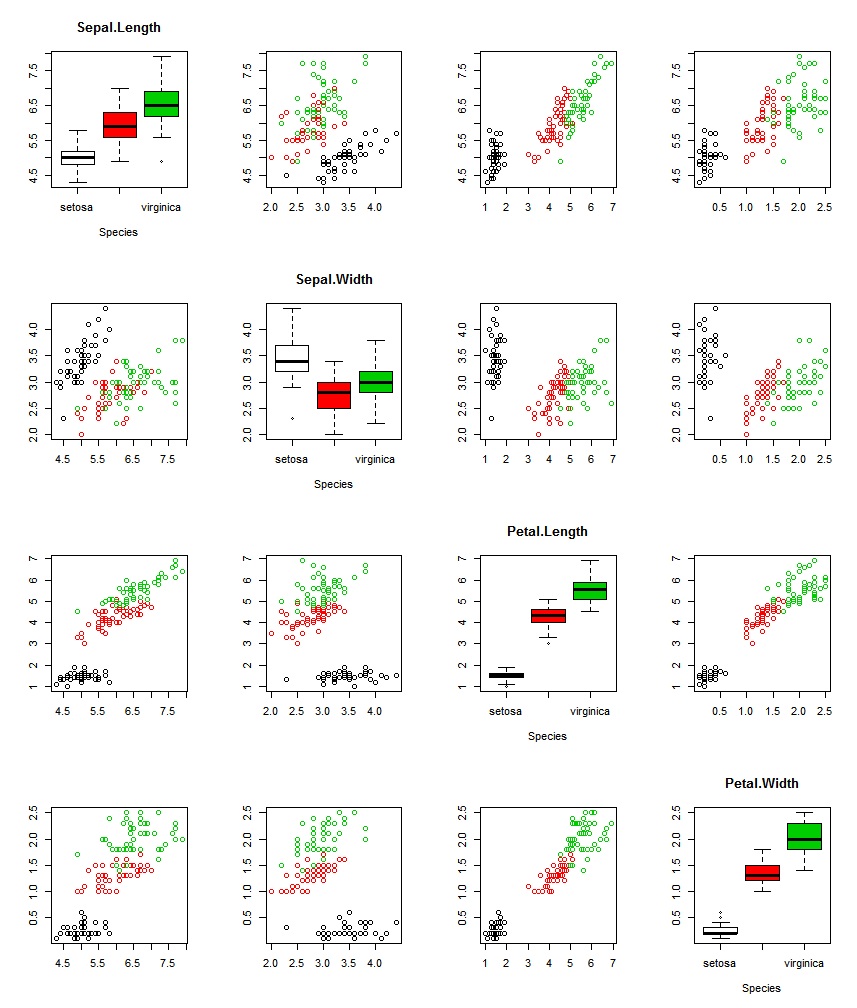
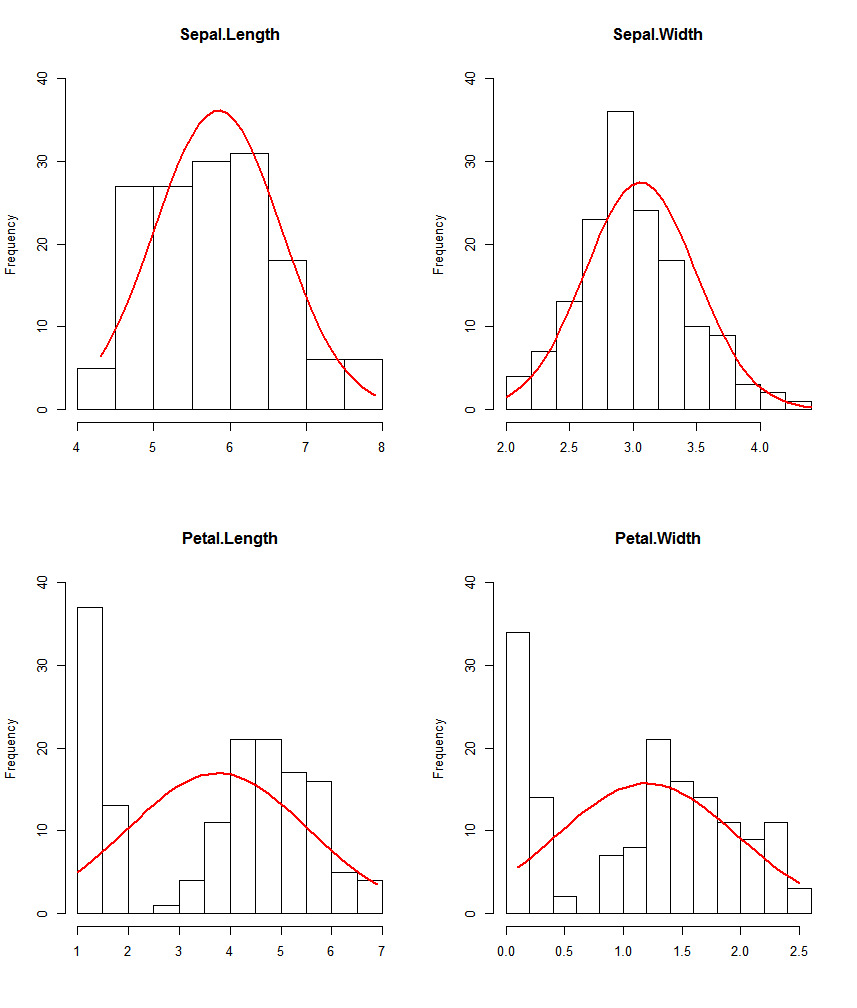
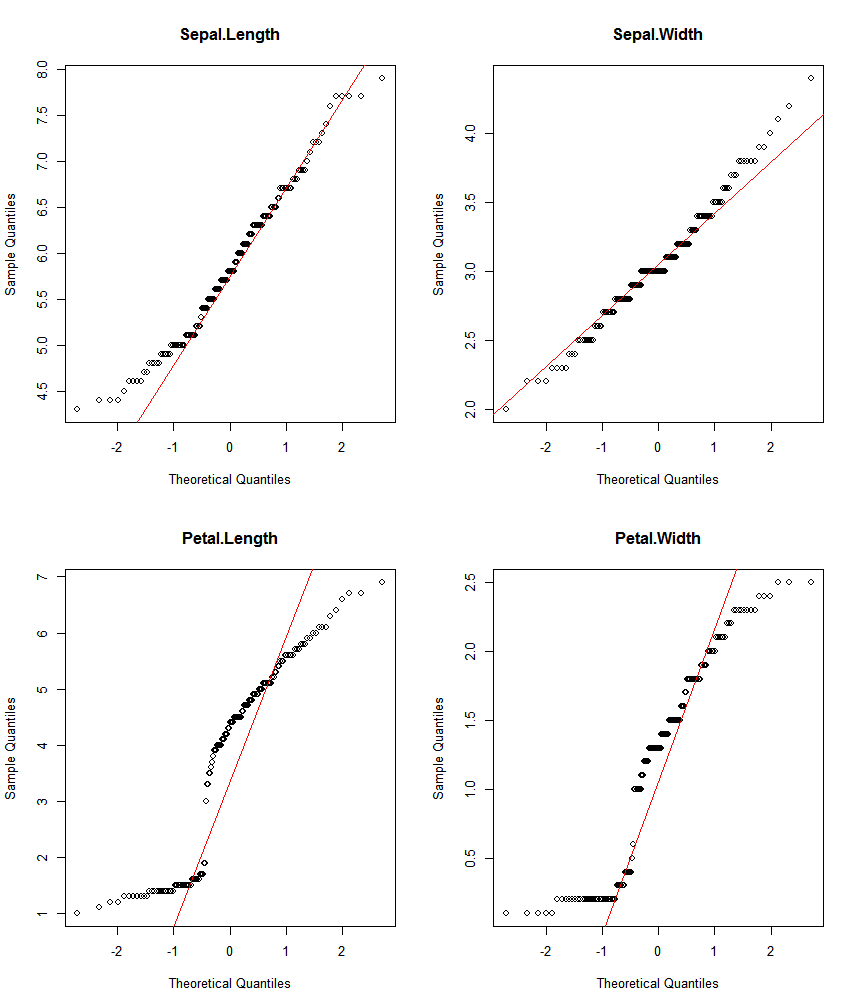
Multivariate Analysis

Homework 1

M052040003 鍾冠毅

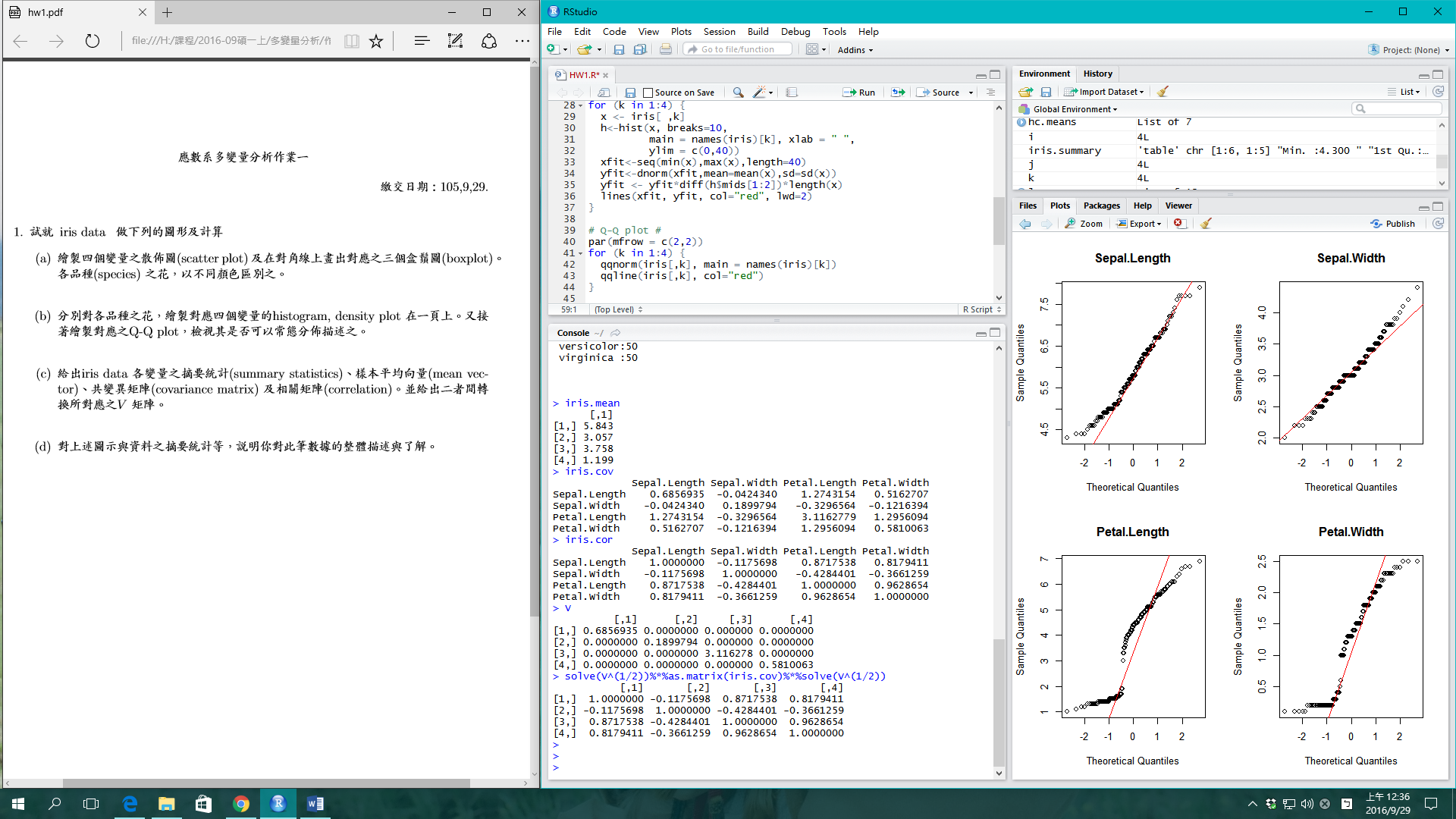
1.a.

1.b.



除了Sepal.Width較接近常態分佈以外，其他三者明顯不像常態分佈。

1.c.



1.d. 由相關係數矩陣可知道Sepal.Length分別與Petal.Length、Petal.Width有高度正相關。另外Petal.Length、Petal.Width兩者間也有高達0.9628654的相關係數。同時觀察分散圖矩陣，亦稍微看得出其相關性。

由直方圖可以看見僅Sepal.Width較符合常態分布，與Q-Q Plot比較也可以應證上述所言，但是要有更嚴謹的常態性分析，可藉由假設檢定之方法，檢定其常態性與否。

**Appendix**

attach(iris)

### 1.a ###

par(mfrow = c(4, 4))

for (i in 1:4) {

for (j in 1:4){

if(i==j){

plot(iris[, i]~iris$Species,

xlab="Species", ylab = " ",

main = names(iris)[i],

col = c(0,2,3))

}

else{

plot(iris[,i]~iris[,j],

xlab = " ", ylab = " "

, col = Species)

}

}

}

### 1.b ###

# histogram and density plot #

par(mfrow = c(2,2))

for (k in 1:4) {

x <- iris[ ,k]

h<-hist(x, breaks=10,

main = names(iris)[k], xlab = " ",

ylim = c(0,40))

xfit<-seq(min(x),max(x),length=40)

yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))

yfit <- yfit\*diff(h$mids[1:2])\*length(x)

lines(xfit, yfit, col="red", lwd=2)

}

# Q-Q plot #

par(mfrow = c(2,2))

for (k in 1:4) {

qqnorm(iris[,k], main = names(iris)[k])

qqline(iris[,k], col="red")

}

### 1.c ###

iris.summary <- summary(iris)

iris.mean <- matrix(data = c(5.843, 3.057, 3.758, 1.199), 4, 1)

iris.cov <- cov(iris[, 1:4])

iris.cor <- cor(iris[, 1:4])

V <- diag(NA, 4)

for (k in 1:4) {

V[k, k] <- cov(iris[, 1:4])[k, k]

}

solve(V^(1/2))%\*%as.matrix(iris.cov)%\*%solve(V^(1/2)