The dataset in the Excel workbook, “Assignment 3a Data”, contains a single outlier – all the other points were generated from a trivariate normal distribution. Use each of the following methods to try to detect the outlier. Comment on which one (if any) seems to work, and why some apparently fail.

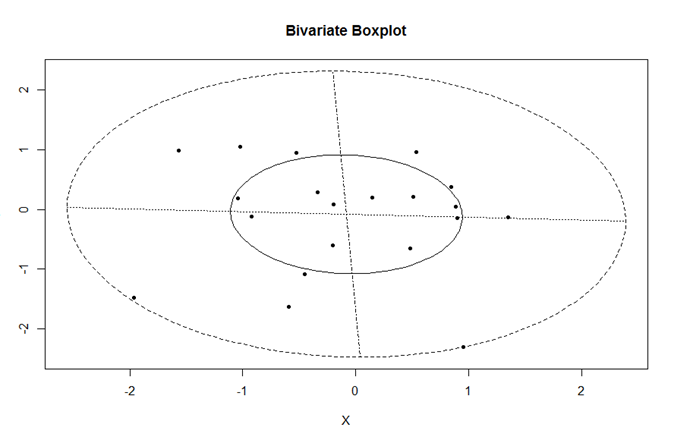
* 1. Univariate boxplots

setwd("C:\\Users\\Kun\\Desktop\\homework")

data<-read.csv("Assignment 3a Data.csv",header=0)

**###the univariate boxplots.**

**boxplot(data, pch=20, main="Univariate Boxplots")**



* 1. Bivariate boxplots

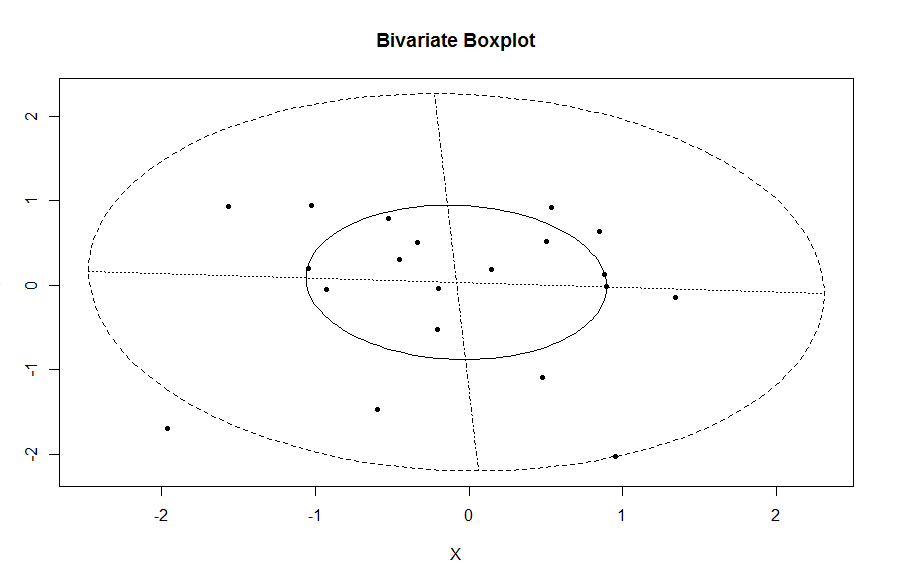
**###Bivariate boxplots**

library("MVA")

X<-data[1]

Y<-data[2]

bvbox(cbind(X, Y),7, pch=20,main="Bivariate Boxplot")



X<-data[1]

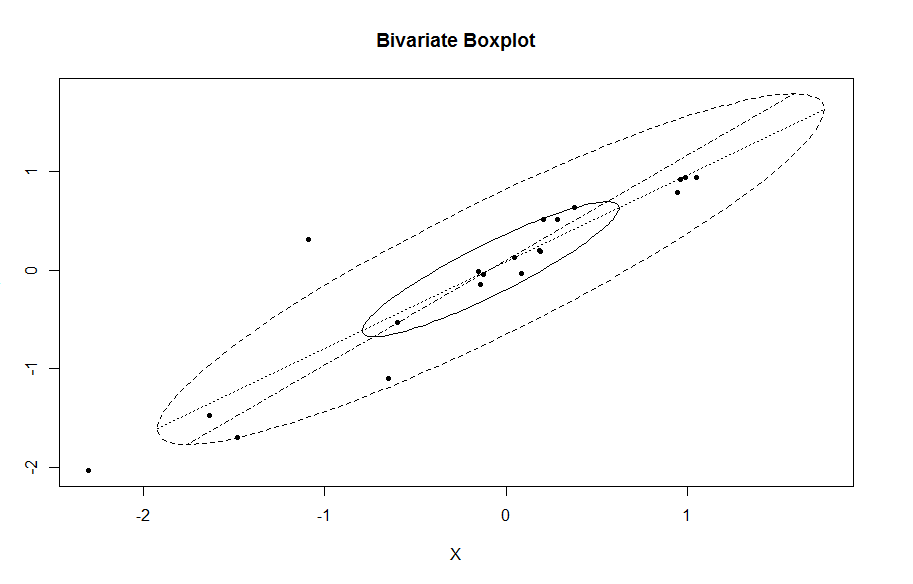
Y<-data[3]

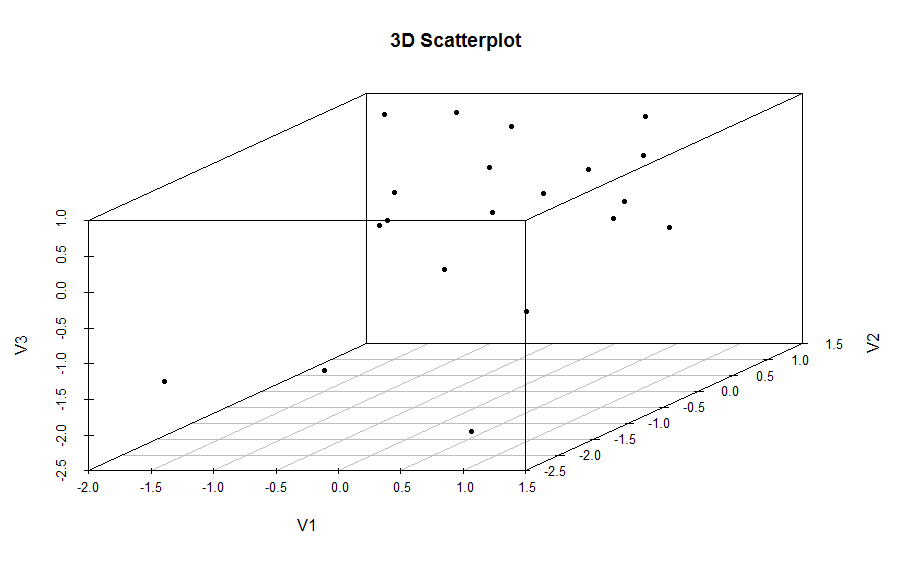
bvbox(cbind(X, Y),7, pch=20,main="Bivariate Boxplot")

X<-data[2]

Y<-data[3]

bvbox(cbind(X, Y),7, pch=20,main="Bivariate Boxplot")



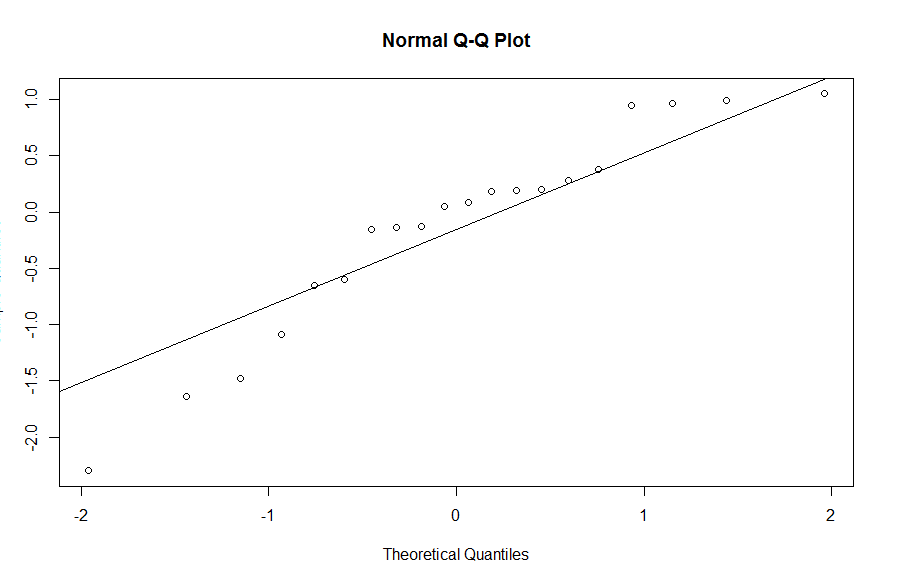
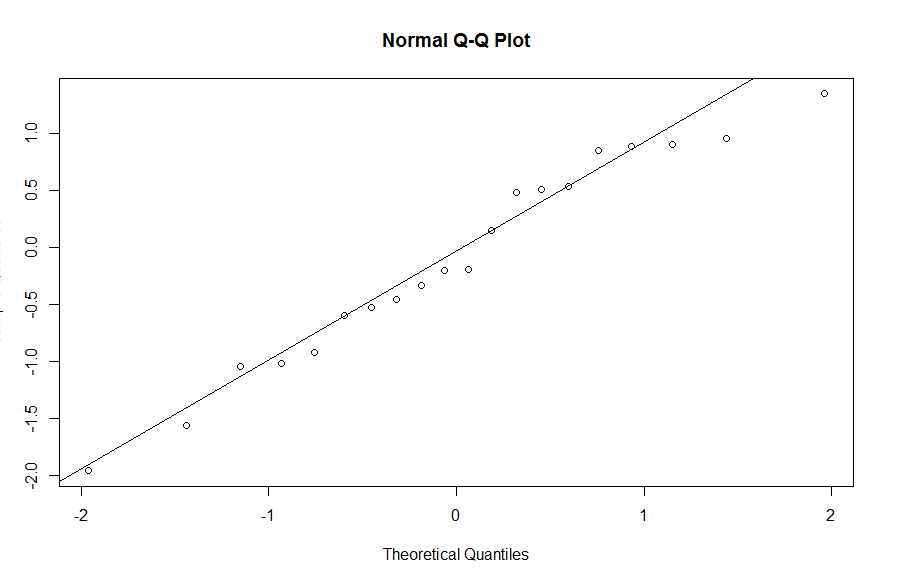
* 1. A three-dimensional scatterplot、

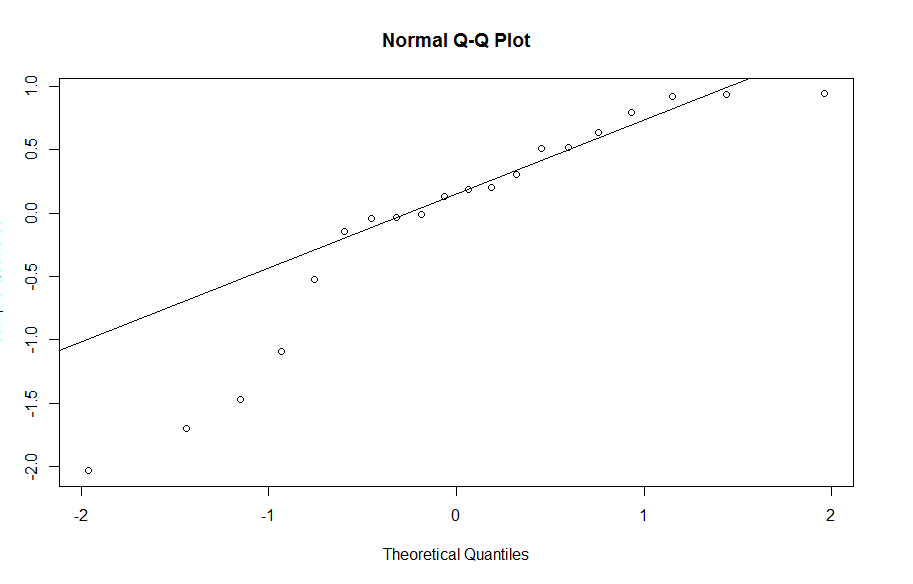
**###three-dimensional scatterplot**

**library(scatterplot3d)**

**scatterplot3d(data,pch=20,main="3D Scatterplot")**

* 1. A chi-square plot.





We find the outlier is (-1.96054, -1.48029,-1.69866). Only the three-dimensional scatterplot can be used to find the outlier.

Univariate boxplots only show single variable. Bivariate boxplots only contain two variables. Chi-square plots cannot combine three variables together in one graph. Therefore, are fail to detect the outliers from a trivariate normal distribution