Homework 1

Jiaqi Li

**A.**

> #A----------------------------------------------------------------------------------

> bodyfat=read.table("BODY\_FAT.TXT", header=T, sep="")

> head(bodyfat)

Density SiriBFperc Over45 Weight Height NeckC ChestC AbdomenC HipC ThighC KneeC AnkleC BicepsC ForearmC WristC

1 1.0708 12.3 0 154.25 67.75 36.2 93.1 85.2 94.5 59.0 37.3 21.9 32.0 27.4 17.1

2 1.0853 6.1 0 173.25 72.25 38.5 93.6 83.0 98.7 58.7 37.3 23.4 30.5 28.9 18.2

3 1.0414 25.3 0 154.00 66.25 34.0 95.8 87.9 99.2 59.6 38.9 24.0 28.8 25.2 16.6

4 1.0751 10.4 0 184.75 72.25 37.4 101.8 86.4 101.2 60.1 37.3 22.8 32.4 29.4 18.2

5 1.0340 28.7 0 184.25 71.25 34.4 97.3 100.0 101.9 63.2 42.2 24.0 32.2 27.7 17.7

6 1.0502 20.9 0 210.25 74.75 39.0 104.5 94.4 107.8 66.0 42.0 25.6 35.7 30.6 18.8

> bodyfat.recompute=(495/bodyfat$Density)-450

> head(round(bodyfat.recompute,digits=1))

[1] 12.3 6.1 25.3 10.4 28.7 21.3

> bodyfat.recompute>100

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[20] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[39] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[58] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[77] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[96] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[115] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[134] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[153] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[172] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[191] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[210] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[248] FALSE FALSE FALSE FALSE FALSE

> bodyfat.recompute<0

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[20] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[39] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[58] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[77] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[96] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[115] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[134] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[153] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[172] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[191] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[210] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[248] FALSE FALSE FALSE FALSE FALSE

>

**Then, there is no value that is smaller than 0 and bigger than 100 in the dataset of bodyfat.recompute.**

> bodyfat$SiriBFperc==round(bodyfat.recompute,digits=1)

[1] TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[20] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[39] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[58] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE FALSE

[77] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[96] FALSE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE

[115] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[134] TRUE TRUE FALSE TRUE TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[153] TRUE TRUE TRUE FALSE TRUE TRUE FALSE TRUE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE

[172] TRUE TRUE FALSE TRUE TRUE TRUE FALSE TRUE FALSE TRUE FALSE TRUE TRUE FALSE TRUE FALSE TRUE TRUE TRUE

[191] TRUE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE FALSE TRUE TRUE TRUE FALSE FALSE FALSE TRUE FALSE TRUE

[210] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE

[229] TRUE FALSE FALSE TRUE TRUE TRUE TRUE TRUE FALSE FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE TRUE TRUE

[248] FALSE TRUE TRUE TRUE TRUE

**There are some erroneous values in the variable “SiriBFPerc”.**

**Thus, employ my recomputed variable.**

> bodyfat[,2]<-round(bodyfat.recompute,digits=1)

> colnames(bodyfat)[colnames(bodyfat)=="SiriBFperc"] <- "bodyfat.percent"

> head(bodyfat)

Density bodyfat.percent Over45 Weight Height NeckC ChestC AbdomenC HipC ThighC KneeC AnkleC BicepsC ForearmC WristC

1 1.0708 12.3 0 154.25 67.75 36.2 93.1 85.2 94.5 59.0 37.3 21.9 32.0 27.4 17.1

2 1.0853 6.1 0 173.25 72.25 38.5 93.6 83.0 98.7 58.7 37.3 23.4 30.5 28.9 18.2

3 1.0414 25.3 0 154.00 66.25 34.0 95.8 87.9 99.2 59.6 38.9 24.0 28.8 25.2 16.6

4 1.0751 10.4 0 184.75 72.25 37.4 101.8 86.4 101.2 60.1 37.3 22.8 32.4 29.4 18.2

5 1.0340 28.7 0 184.25 71.25 34.4 97.3 100.0 101.9 63.2 42.2 24.0 32.2 27.7 17.7

6 1.0502 21.3 0 210.25 74.75 39.0 104.5 94.4 107.8 66.0 42.0 25.6 35.7 30.6 18.8

**Check also variables “Weight”, “Height”, “AbdomenC”**

> weight.dataset=bodyfat[,4]

> height.dataset=bodyfat[,5]

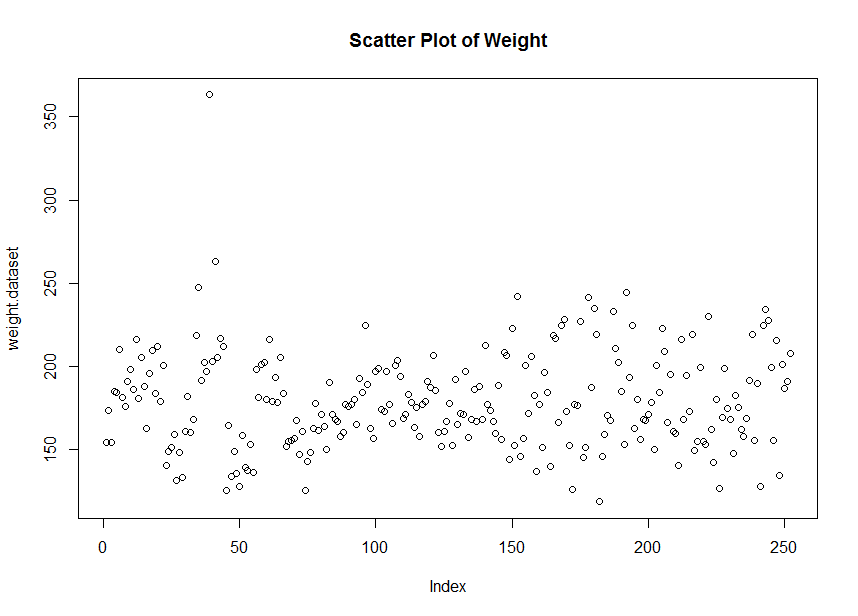
> AbdomenC.dataset=bodyfat[,8]

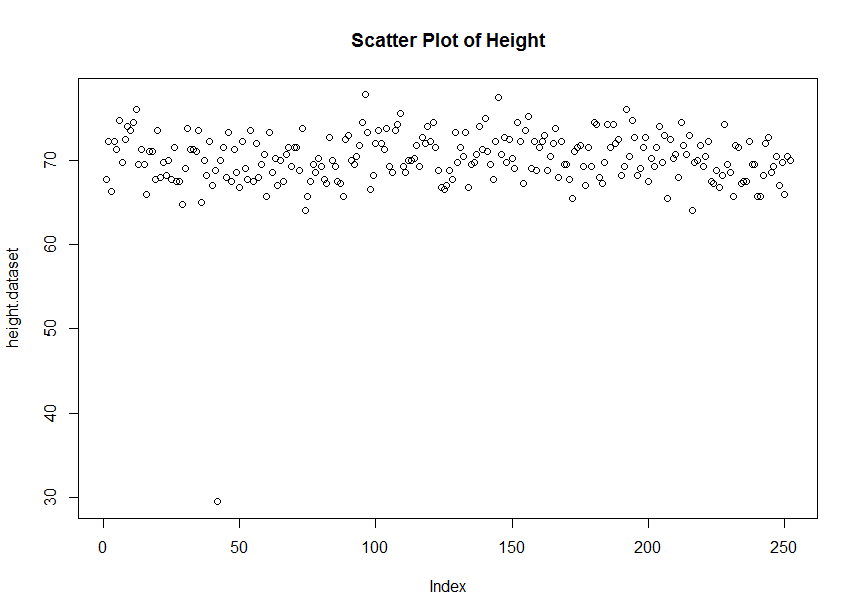
>

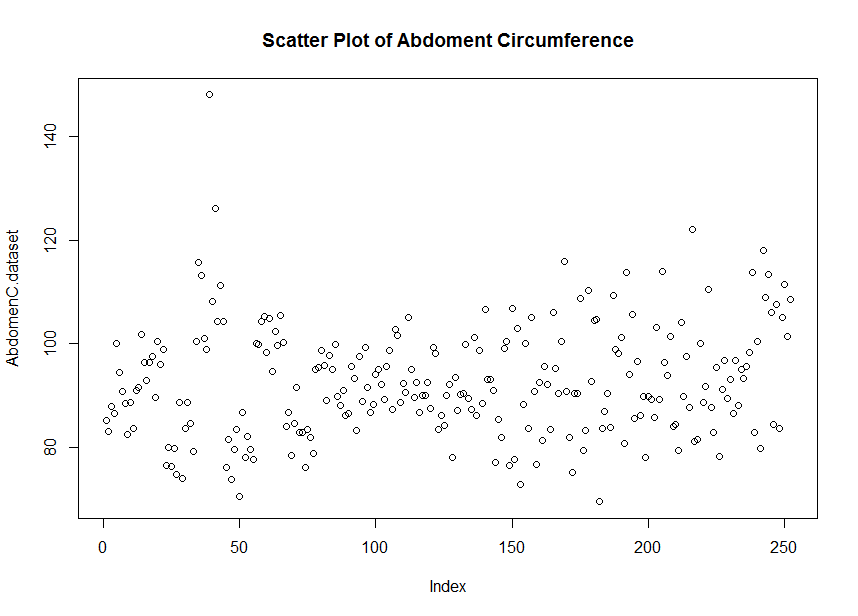
> plot(weight.dataset,main="Scatter Plot of Weight")

> plot(height.dataset,main="Scatter Plot of Height")

> plot(AbdomenC.dataset,main="Scatter Plot of Abdoment Circumference")







**By observing the scatter plots, we can see some obvious mistakes.**

**Thus, we want to find these mistakes and remove them.**

> bodyfat$Weight > 350

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[20] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[39] TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[58] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[77] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[96] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[115] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[134] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[153] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[172] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[191] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[210] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[248] FALSE FALSE FALSE FALSE FALSE

> bodyfat$AbdomenC > 140

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[20] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[39] TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[58] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[77] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[96] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[115] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[134] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[153] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[172] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[191] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[210] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[248] FALSE FALSE FALSE FALSE FALSE

> bodyfat<-bodyfat[-39,]

>

> bodyfat$Height < 40

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[20] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[39] FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[58] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[77] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[96] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[115] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[134] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[153] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[172] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[191] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[210] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[248] FALSE FALSE FALSE FALSE

> bodyfat<-bodyfat[-41,]

**Now, we should have a good dataset for further analysis.**

> weight.dataset=bodyfat[,4]

> height.dataset=bodyfat[,5]

> AbdomenC.dataset=bodyfat[,8]

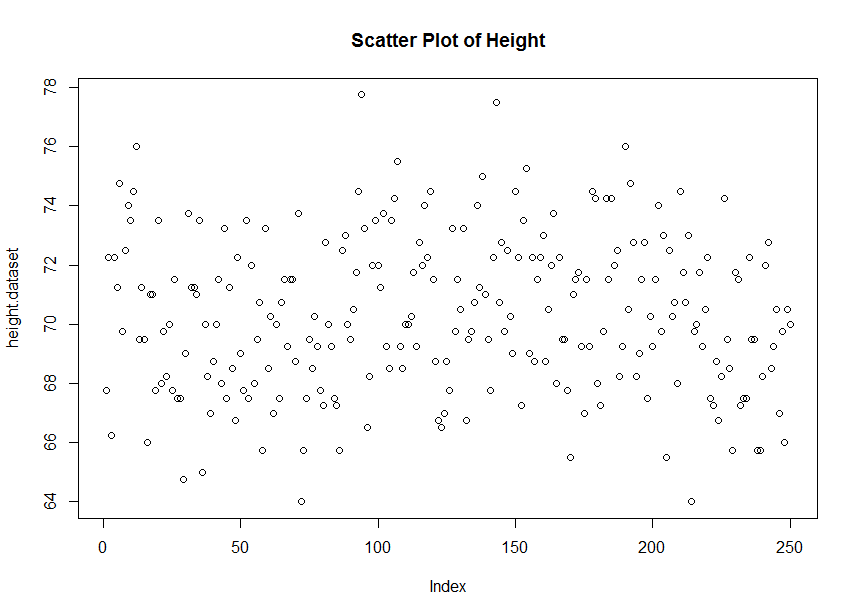
>

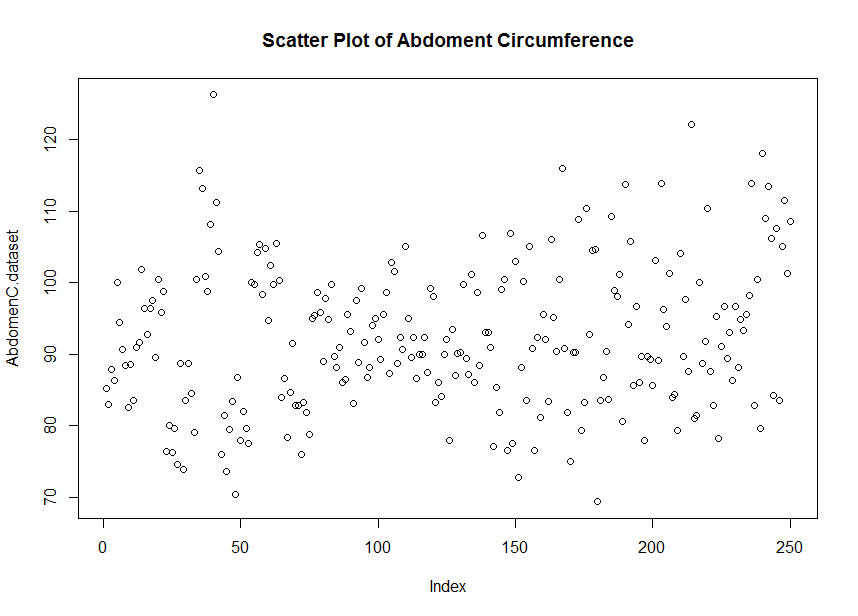
> plot(weight.dataset,main="Scatter Plot of Weight")

> plot(height.dataset,main="Scatter Plot of Height")

> plot(AbdomenC.dataset,main="Scatter Plot of Abdoment Circumference")







**B.**

> #B--------------------------------------------------------------------------------

> summary(bodyfat$bodyfat.percent)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-3.60 12.40 19.20 18.97 25.18 47.50

> sd(bodyfat$bodyfat.percent)

[1] 8.393672

> IQR(bodyfat$bodyfat.percent)

[1] 12.775

> hist(bodyfat$bodyfat.percent)

> boxplot(bodyfat$bodyfat.percent,

+ main="Boxplot of Bodyfat Percentage")

>

> summary(bodyfat$Weight)

Min. 1st Qu. Median Mean 3rd Qu. Max.

118.5 158.5 176.1 178.1 196.8 262.8

> sd(bodyfat$Weight)

[1] 27.03549

> IQR(bodyfat$Weight)

[1] 38.25

> hist(bodyfat$Weight)

> boxplot(bodyfat$Weight,

+ main="Boxplot of Weight")

>

> summary(bodyfat$Height)

Min. 1st Qu. Median Mean 3rd Qu. Max.

64.00 68.25 70.00 70.30 72.25 77.75

> sd(bodyfat$Height)

[1] 2.616644

> IQR(bodyfat$Height)

[1] 4

> hist(bodyfat$Height)

> boxplot(bodyfat$Height,

+ main="Boxplot of Height")

>

> summary(bodyfat$AbdomenC)

Min. 1st Qu. Median Mean 3rd Qu. Max.

69.40 84.53 90.90 92.29 99.17 126.20

> sd(bodyfat$AbdomenC)

[1] 10.20744

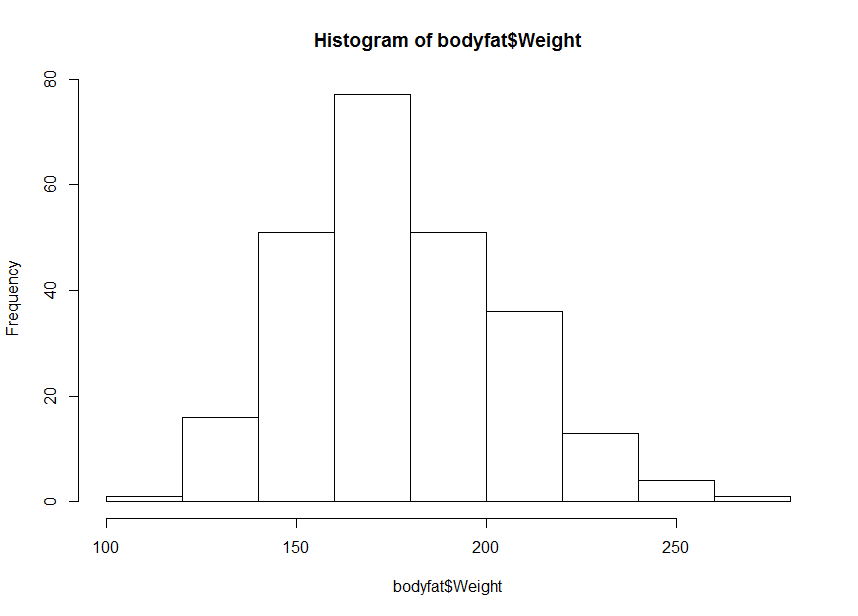
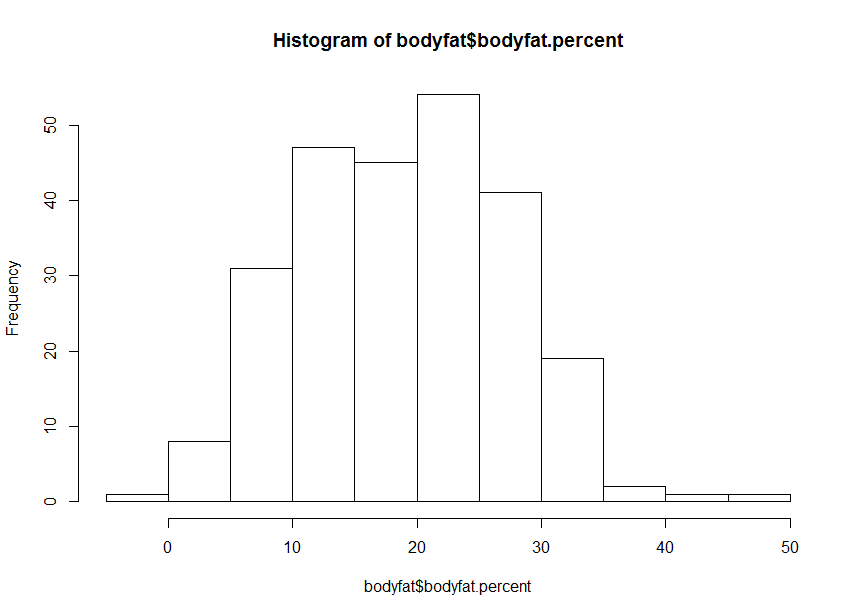
> IQR(bodyfat$AbdomenC)

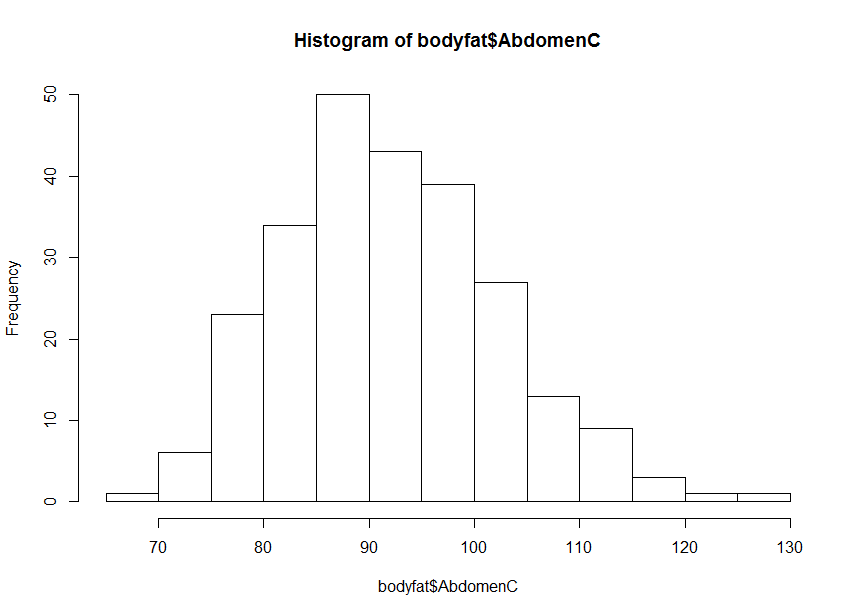
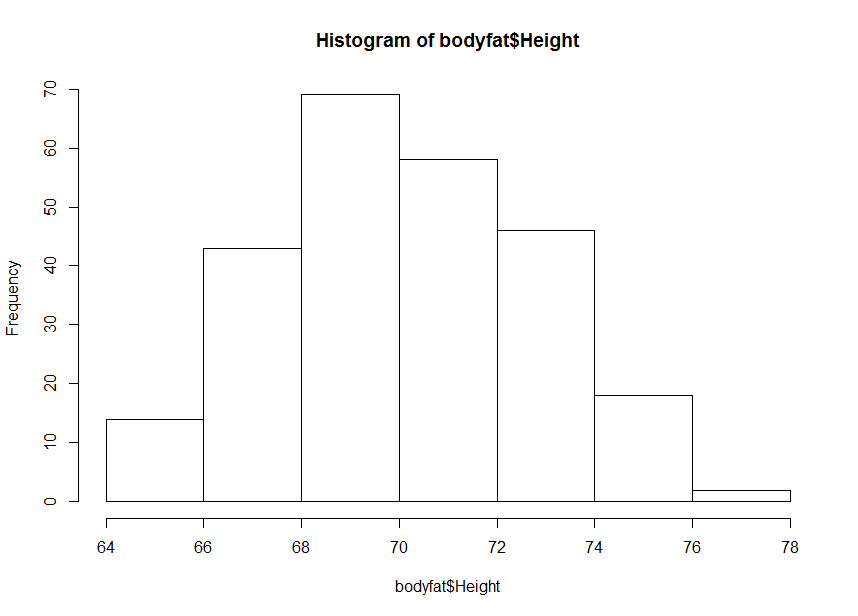
[1] 14.65

> hist(bodyfat$AbdomenC)

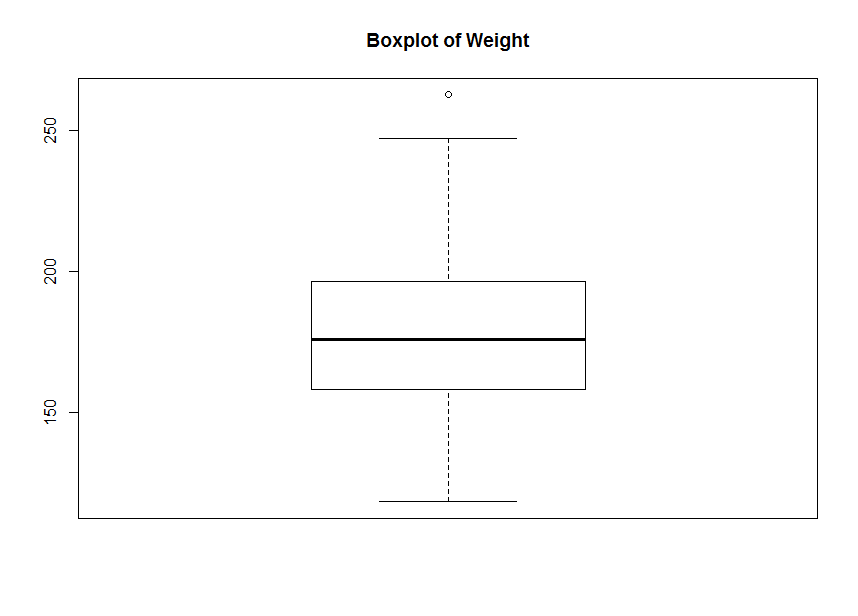
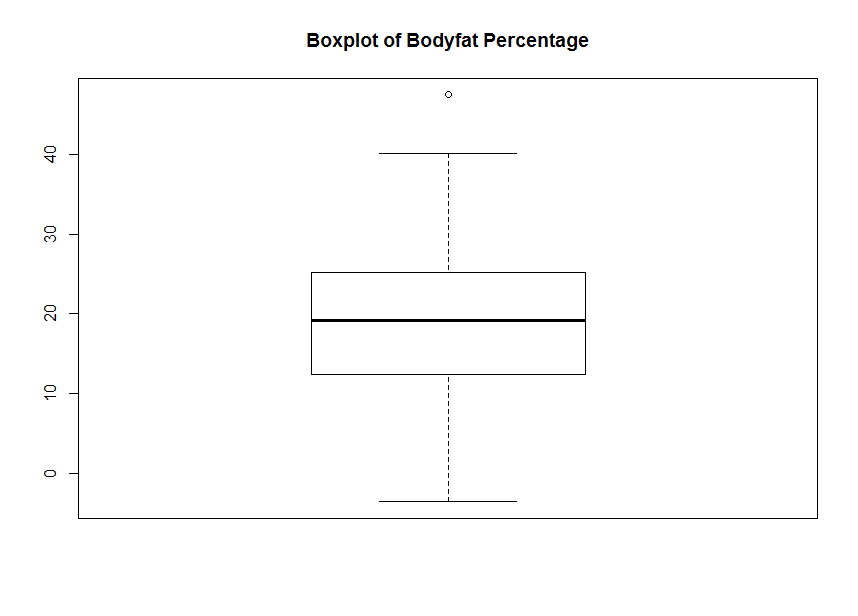
> boxplot(bodyfat$AbdomenC,

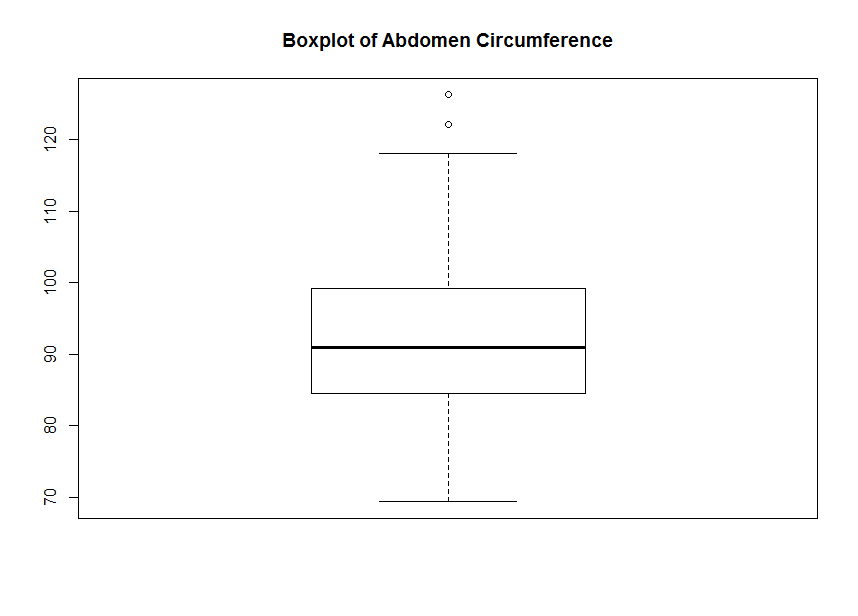
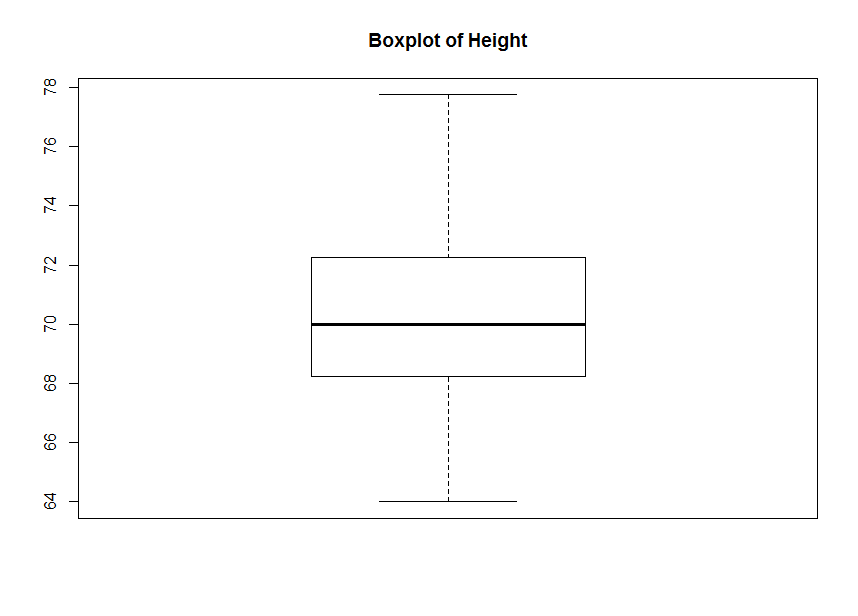
+ main="Boxplot of Abdomen Circumference")





**All these distributions appear to be symmetric and bell-shaped.**





**Based on the boxplots, there are a few extreme values in distributions of Bodyfat Percentage, Weight, Abdomen Circumference.**

**Hypothesis test on body fat percentage:**

**H0 : average body fat percentage ≥20%**

**H1 : average body fat percentage < 20%**

> bodyfat.percent.test=as.matrix(bodyfat$bodyfat.percent)

> z.bodyfat.percent=(mean(bodyfat.percent.test)-20)/(sd(bodyfat.percent.test)/sqrt(250))

> z.bodyfat.percent

[1] -1.938732

> pvalue.bodyfat.percent=1-pnorm(z.bodyfat.percent)

> pvalue.bodyfat.percent

[1] 0.973733

**Thus, p-value = 0.973733 > α , reject H0, which means the average body fat percentage does not exceed 20%**

**Hypothesis test on weight:**

**H0 : average weight ≥ 180 pounds**

**H1 : average weight < 180 pounds**

> weight.test=as.matrix(bodyfat$Weight)

> z.weight=(mean(weight.test)-180)/(sd(weight.test)/sqrt(250))

> z.weight

[1] -1.121018

> pvalue.weight=1-pnorm(z.weight)

> pvalue.weight

[1] 0.8688599

**Thus, p-value =0.8688599 > α ,reject H0, which means the average weight does not exceed 180 pounds.**

**C.**

> #C-------------------------------------------------------------------------------

> attach(bodyfat)

The following objects are masked from bodyfat (pos = 3):

AbdomenC, AnkleC, BicepsC, bodyfat.percent, ChestC, Density, ForearmC, Height, HipC, KneeC, NeckC,

Over45, ThighC, Weight, WristC

The following objects are masked from bodyfat (pos = 4):

AbdomenC, AnkleC, BicepsC, bodyfat.percent, ChestC, Density, ForearmC, Height, HipC, KneeC, NeckC,

Over45, ThighC, Weight, WristC

The following objects are masked from bodyfat (pos = 5):

AbdomenC, AnkleC, BicepsC, bodyfat.percent, ChestC, Density, ForearmC, Height, HipC, KneeC, NeckC,

Over45, ThighC, Weight, WristC

The following objects are masked from bodyfat (pos = 6):

AbdomenC, AnkleC, BicepsC, bodyfat.percent, ChestC, Density, ForearmC, Height, HipC, KneeC, NeckC,

Over45, ThighC, Weight, WristC

>

> cor(bodyfat.percent,Weight)

[1] 0.5993419

> model.weight=lm(bodyfat.percent~Weight)

> summary(model.weight)

Call:

lm(formula = bodyfat.percent ~ Weight)

Residuals:

Min 1Q Median 3Q Max

-27.208 -4.597 0.057 4.981 20.916

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -14.16636 2.84250 -4.984 1.17e-06 \*\*\*

Weight 0.18608 0.01578 11.791 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.733 on 248 degrees of freedom

Multiple R-squared: 0.3592, Adjusted R-squared: 0.3566

F-statistic: 139 on 1 and 248 DF, p-value: < 2.2e-16

> sigma2.weight=sum((model.weight$residuals)^2)/248

> sigma2.weight

[1] 45.32803

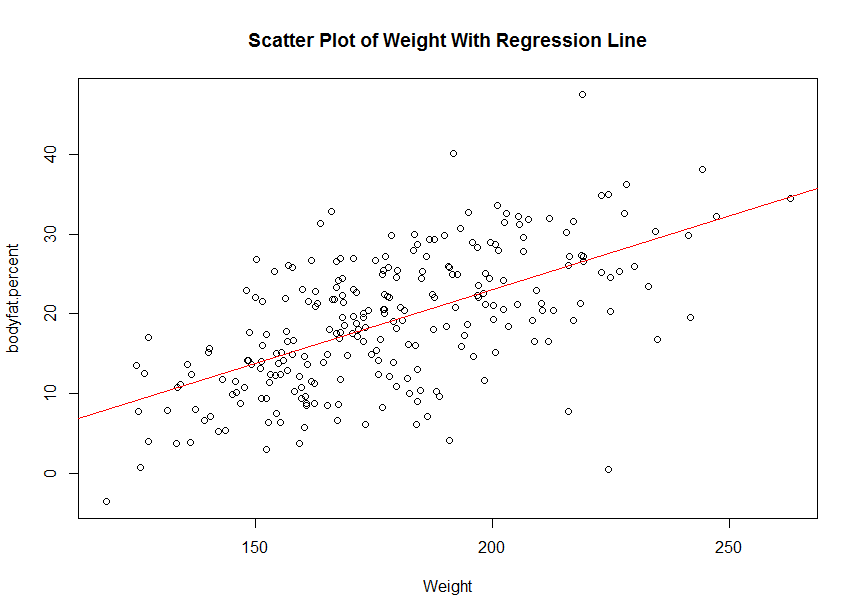
> beta0.weight=-14.16636

> beta1.weight=0.18608

> plot(Weight,bodyfat.percent,

+ main="Scatter Plot of Weight With Regression Line")

> abline(beta0.weight,beta1.weight,col="red")



> cor(bodyfat.percent,Height)

[1] -0.04682211

> model.height=lm(bodyfat.percent~Height)

> summary(model.height)

Call:

lm(formula = bodyfat.percent ~ Height)

Residuals:

Min 1Q Median 3Q Max

-22.9167 -6.5501 0.3023 6.2215 27.5825

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 29.5300 14.3146 2.063 0.0402 \*

Height -0.1502 0.2035 -0.738 0.4611

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.401 on 248 degrees of freedom

Multiple R-squared: 0.002192, Adjusted R-squared: -0.001831

F-statistic: 0.5449 on 1 and 248 DF, p-value: 0.4611

> sigma2.height=sum((model.height$residuals)^2)/248

> sigma2.height

[1] 70.58273

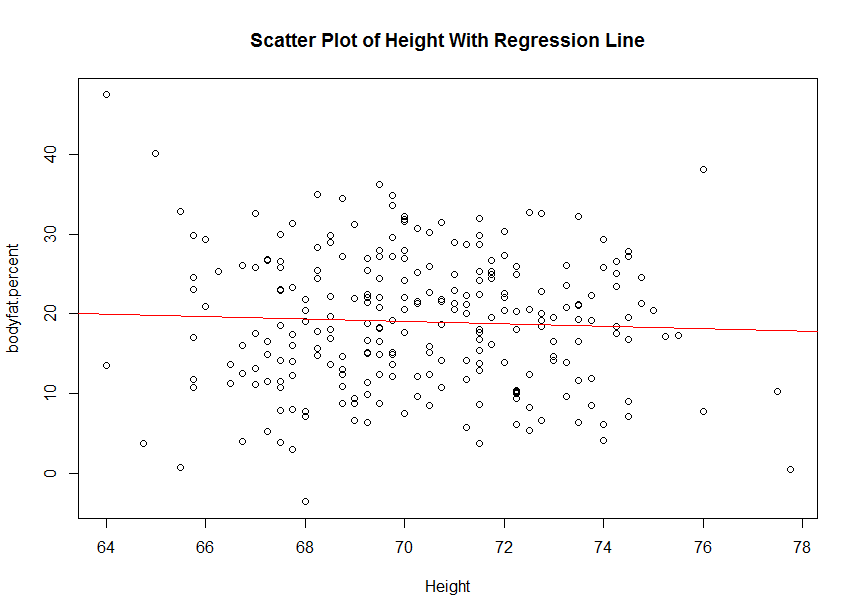
> beta0.height=29.5300

> beta1.height=-0.150

> plot(Height,bodyfat.percent,

+ main="Scatter Plot of Height With Regression Line")

> abline(beta0.height,beta1.height,col="red")



> cor(bodyfat.percent,AbdomenC)

[1] 0.811426

> model.abdomenc=lm(bodyfat.percent~AbdomenC)

> summary(model.abdomenc)

Call:

lm(formula = bodyfat.percent ~ AbdomenC)

Residuals:

Min 1Q Median 3Q Max

-23.1836 -3.5224 0.2425 3.1878 12.8528

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -42.60690 2.83354 -15.04 <2e-16 \*\*\*

AbdomenC 0.66724 0.03052 21.86 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.916 on 248 degrees of freedom

Multiple R-squared: 0.6584, Adjusted R-squared: 0.657

F-statistic: 478 on 1 and 248 DF, p-value: < 2.2e-16

> sigma2.abdomenc=sum((model.abdomenc$residuals)^2)/248

> sigma2.abdomenc

[1] 24.16318

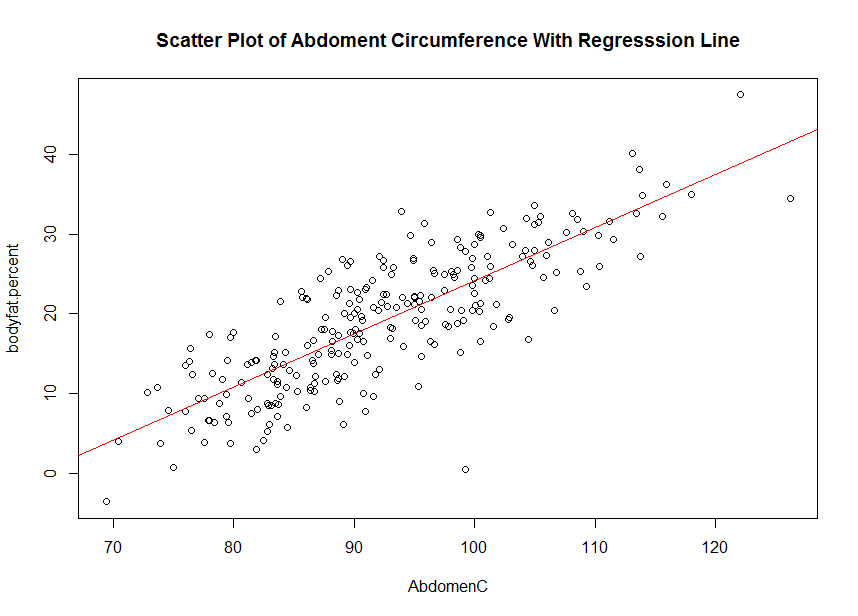
> beta0.abdomenc=-42.60690

> beta1.abdomenc=0.66724

> plot(AbdomenC,bodyfat.percent,

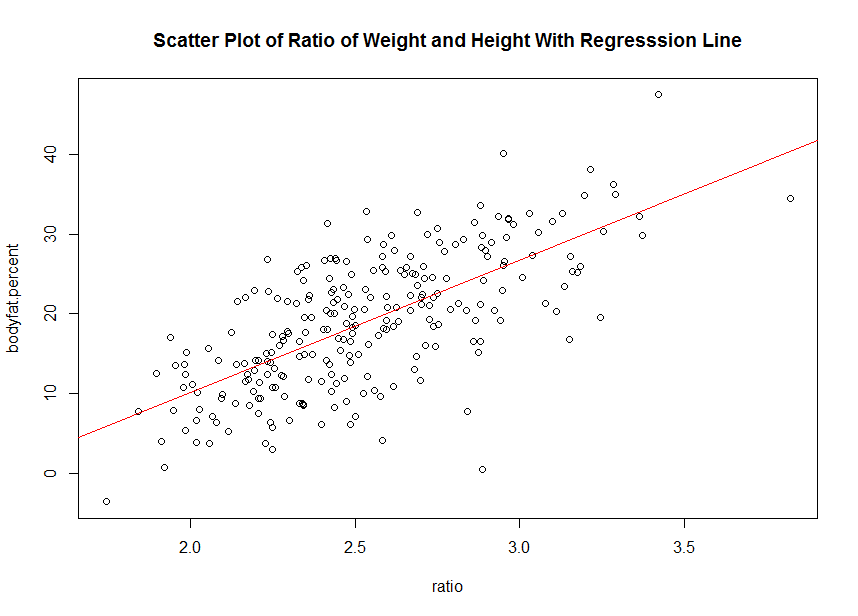
+ main="Scatter Plot of Abdoment Circumference With Regresssion Line")

> abline(beta0.abdomenc,beta1.abdomenc,col="red")



**D.**

|  |
| --- |
| > #D-------------------------------------------------------------------------------  > ratio=Weight/Height  > cor(bodyfat.percent,ratio)  [1] 0.6862212  > model.ratio=lm(bodyfat.percent~ratio)  > summary(model.ratio)  Call:  lm(formula = bodyfat.percent ~ ratio)  Residuals:  Min 1Q Median 3Q Max  -24.5140 -4.0119 0.1253 4.2037 14.2930  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -22.999 2.851 -8.066 3.11e-14 \*\*\*  ratio 16.594 1.117 14.857 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 6.118 on 248 degrees of freedom  Multiple R-squared: 0.4709, Adjusted R-squared: 0.4688  F-statistic: 220.7 on 1 and 248 DF, p-value: < 2.2e-16  > sigma2.ratio=sum((model.ratio$residuals)^2)/248  > sigma2.ratio  [1] 37.4274  > beta0.ratio=-22.999  > beta1.ratio=16.594  > plot(ratio,bodyfat.percent,  + main="Scatter Plot of Ratio of Weight and Height With Regresssion Line")  > abline(beta0.ratio,beta1.ratio,col="red") |
|  |
|  |



**E.**

> #E-------------------------------------------------------------------------------

> cor(ratio,AbdomenC)

[1] 0.9236815

> model.rA=lm(ratio~AbdomenC)

> summary(model.rA)

Call:

lm(formula = ratio ~ AbdomenC)

Residuals:

Min 1Q Median 3Q Max

-0.34024 -0.07587 -0.00816 0.07772 0.38230

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.3695357 0.0768221 -4.81 2.62e-06 \*\*\*

AbdomenC 0.0314111 0.0008274 37.96 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1333 on 248 degrees of freedom

Multiple R-squared: 0.8532, Adjusted R-squared: 0.8526

F-statistic: 1441 on 1 and 248 DF, p-value: < 2.2e-16

> sigma2.rA=sum((model.rA$residuals)^2)/248

> sigma2.rA

[1] 0.01776096

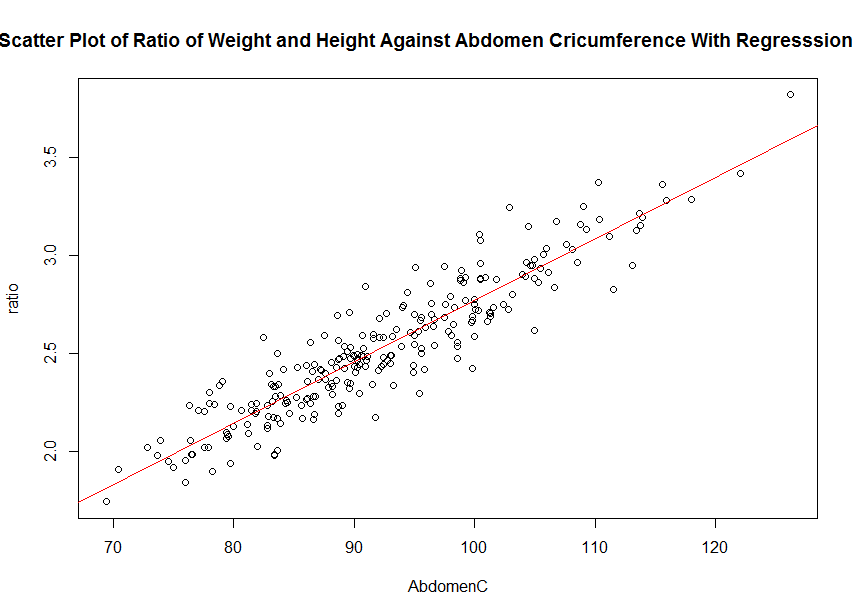
> beta0.rA=-0.3695357

> beta1.rA=0.0314111

> plot(AbdomenC,ratio,

+ main="Scatter Plot of Ratio of Weight and Height Against Abdomen Cricumference With Regresssion Line")

> abline(beta0.rA,beta1.rA,col="red")



**R code:**

**setwd("//udrive.win.psu.edu/Users/j/q/jql5883/Desktop/math462")**

**getwd()**

**#A----------------------------------------------------------------------------------**

**bodyfat=read.table("BODY\_FAT.TXT", header=T, sep="")**

**head(bodyfat)**

**bodyfat.recompute=(495/bodyfat$Density)-450**

**head(round(bodyfat.recompute,digits=1))**

**bodyfat.recompute>100**

**bodyfat.recompute<0**

**bodyfat$SiriBFperc==round(bodyfat.recompute,digits=1)**

**bodyfat[,2]<-round(bodyfat.recompute,digits=1)**

**colnames(bodyfat)[colnames(bodyfat)=="SiriBFperc"] <- "bodyfat.percent"**

**head(bodyfat)**

**weight.dataset=bodyfat[,4]**

**height.dataset=bodyfat[,5]**

**AbdomenC.dataset=bodyfat[,8]**

**plot(weight.dataset,main="Scatter Plot of Weight")**

**plot(height.dataset,main="Scatter Plot of Height")**

**plot(AbdomenC.dataset,main="Scatter Plot of Abdoment Circumference")**

**bodyfat$Weight > 350**

**bodyfat$AbdomenC > 140**

**bodyfat<-bodyfat[-39,]**

**bodyfat$Height < 40**

**bodyfat<-bodyfat[-41,]**

**weight.dataset=bodyfat[,4]**

**height.dataset=bodyfat[,5]**

**AbdomenC.dataset=bodyfat[,8]**

**plot(weight.dataset,main="Scatter Plot of Weight")**

**plot(height.dataset,main="Scatter Plot of Height")**

**plot(AbdomenC.dataset,main="Scatter Plot of Abdoment Circumference")**

**#B--------------------------------------------------------------------------------**

**summary(bodyfat$bodyfat.percent)**

**sd(bodyfat$bodyfat.percent)**

**IQR(bodyfat$bodyfat.percent)**

**hist(bodyfat$bodyfat.percent)**

**boxplot(bodyfat$bodyfat.percent,**

**main="Boxplot of Bodyfat Percentage")**

**summary(bodyfat$Weight)**

**sd(bodyfat$Weight)**

**IQR(bodyfat$Weight)**

**hist(bodyfat$Weight)**

**boxplot(bodyfat$Weight,**

**main="Boxplot of Weight")**

**summary(bodyfat$Height)**

**sd(bodyfat$Height)**

**IQR(bodyfat$Height)**

**hist(bodyfat$Height)**

**boxplot(bodyfat$Height,**

**main="Boxplot of Height")**

**summary(bodyfat$AbdomenC)**

**sd(bodyfat$AbdomenC)**

**IQR(bodyfat$AbdomenC)**

**hist(bodyfat$AbdomenC)**

**boxplot(bodyfat$AbdomenC,**

**main="Boxplot of Abdomen Circumference")**

**bodyfat.percent.test=as.matrix(bodyfat$bodyfat.percent)**

**z.bodyfat.percent=(mean(bodyfat.percent.test)-20)/(sd(bodyfat.percent.test)/sqrt(250))**

**z.bodyfat.percent**

**pvalue.bodyfat.percent=1-pnorm(z.bodyfat.percent)**

**pvalue.bodyfat.percent**

**weight.test=as.matrix(bodyfat$Weight)**

**z.weight=(mean(weight.test)-180)/(sd(weight.test)/sqrt(250))**

**z.weight**

**pvalue.weight=1-pnorm(z.weight)**

**pvalue.weight**

**#C-------------------------------------------------------------------------------**

**attach(bodyfat)**

**cor(bodyfat.percent,Weight)**

**model.weight=lm(bodyfat.percent~Weight)**

**summary(model.weight)**

**sigma2.weight=sum((model.weight$residuals)^2)/248**

**sigma2.weight**

**beta0.weight=-14.16636**

**beta1.weight=0.18608**

**plot(Weight,bodyfat.percent,**

**main="Scatter Plot of Weight With Regression Line")**

**abline(beta0.weight,beta1.weight,col="red")**

**cor(bodyfat.percent,Height)**

**model.height=lm(bodyfat.percent~Height)**

**summary(model.height)**

**sigma2.height=sum((model.height$residuals)^2)/248**

**sigma2.height**

**beta0.height=29.5300**

**beta1.height=-0.150**

**plot(Height,bodyfat.percent,**

**main="Scatter Plot of Height With Regression Line")**

**abline(beta0.height,beta1.height,col="red")**

**cor(bodyfat.percent,AbdomenC)**

**model.abdomenc=lm(bodyfat.percent~AbdomenC)**

**summary(model.abdomenc)**

**sigma2.abdomenc=sum((model.abdomenc$residuals)^2)/248**

**sigma2.abdomenc**

**beta0.abdomenc=-42.60690**

**beta1.abdomenc=0.66724**

**plot(AbdomenC,bodyfat.percent,**

**main="Scatter Plot of Abdoment Circumference With Regresssion Line")**

**abline(beta0.abdomenc,beta1.abdomenc,col="red")**

**#D-------------------------------------------------------------------------------**

**ratio=Weight/Height**

**cor(bodyfat.percent,ratio)**

**model.ratio=lm(bodyfat.percent~ratio)**

**summary(model.ratio)**

**sigma2.ratio=sum((model.ratio$residuals)^2)/248**

**sigma2.ratio**

**beta0.ratio=-22.999**

**beta1.ratio=16.594**

**plot(ratio,bodyfat.percent,**

**main="Scatter Plot of Ratio of Weight and Height With Regresssion Line")**

**abline(beta0.ratio,beta1.ratio,col="red")**

**#E-------------------------------------------------------------------------------**

**cor(ratio,AbdomenC)**

**model.rA=lm(ratio~AbdomenC)**

**summary(model.rA)**

**sigma2.rA=sum((model.rA$residuals)^2)/248**

**sigma2.rA**

**beta0.rA=-0.3695357**

**beta1.rA=0.0314111**

**plot(AbdomenC,ratio,**

**main="Scatter Plot of Ratio of Weight and Height Against Abdomen Cricumference With Regresssion Line")**

**abline(beta0.rA,beta1.rA,col="red")**