**11.42**

a.)

> library(readxl)

> pcb <- read\_excel("~/Downloads/pcb.xls",

+ sheet = "\_A2")

> View(pcb)

>summary(pcb$pcb)

>summary(pcb$pcb52)

>summary(pcb$pcb118)

>summary(pcb$pcb138)

>summary(pcb$pcb180)

> boxplot(pcb$pcb, main="PCB")

> boxplot(pcb$pcb52, main="PCB52")

> boxplot(pcb$pcb118, main="PCB118")

> boxplot(pcb$pcb138, main="PCB138")

> boxplot(pcb$pcb180, main="PCB180")

b.)

> cor(pcb$pcb, pcb$pcb52)

>cor(pcb$pcb, pcb$pcb118)

>cor(pcb$pcb, pcb$pcb138)

>cor(pcb$pcb, pcb$pcb180)

>cor(pcb$pcb52, pcb$pcb118)

>cor(pcb$pcb52, pcb$pcb138)

>cor(pcb$pcb52, pcb$pcb180)

>cor(pcb$pcb118, pcb$pcb138)

>cor(pcb$pcb118, pcb$pcb180)

>cor(pcb$pcb138, pcb$pcb180)

>plot(pcb$pcb52, pcb$pcb, main="PCB vs PCB52", xlab="PCB52", ylab="PCB")

>plot(pcb$pcb118, pcb$pcb, main="PCB vs PCB118", xlab="PCB118", ylab="PCB")

>plot(pcb$pcb138, pcb$pcb, main="PCB vs PCB138", xlab="PCB138", ylab="PCB")

>plot(pcb$pcb180, pcb$pcb, main="PCB vs PCB180", xlab="PCB180", ylab="PCB")

>plot(pcb$pcb118, pcb$pcb52, main="PCB52 vs PCB118", xlab="PCB118", ylab="PCB52")

>plot(pcb$pcb138, pcb$pcb52, main="PCB52 vs PCB138", xlab="PCB138", ylab="PCB52")

>plot(pcb$pcb180, pcb$pcb52, main="PCB52 vs PCB180", xlab="PCB180", ylab="PCB52")

>plot(pcb$pcb138, pcb$pcb118, main="PCB118 vs PCB138", xlab="PCB138", ylab="PCB118")

>plot(pcb$pcb180, pcb$pcb118, main="PCB118 vs PCB180", xlab="PCB180", ylab="PCB118")

>plot(pcb$pcb180, pcb$pcb138, main="PCB138 vs PCB180", xlab="PCB180", ylab="PCB138")

**11.43**

b.)

> subpcb=subset(pcb, select=c("pcb", "pcb52", "pcb118", "pcb138", "pcb180"))

> lm1=lm(pcb~pcb52+pcb118+pcb138+pcb180,data=subpcb)

> summary(lm1)

Call:

lm(formula = pcb ~ pcb52 + pcb118 + pcb138 + pcb180, data = subpcb)

Residuals:

Min 1Q Median 3Q Max

-22.0864 -2.4554 0.0278 2.7726 22.5487

Coefficients:

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

(Intercept) 0.9369 1.2293 0.762 0.449

pcb52 11.8727 0.7290 16.287 < 2e-16 \*\*\*

pcb118 3.7611 0.6424 5.855 1.79e-07 \*\*\*

pcb138 3.8842 0.4978 7.803 7.19e-11 \*\*\*

pcb180 4.1823 0.4318 9.687 3.64e-14 \*\*\*

---

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

Residual standard error: 6.382 on 64 degrees of freedom

Multiple R-squared: 0.9891, Adjusted R-squared: 0.9885

F-statistic: 1456 on 4 and 64 DF, p-value: < 2.2e-16

> anova(lm1)

Analysis of Variance Table

Response: pcb

Df Sum Sq Mean Sq F value Pr(>F)

pcb52 1 85302 85302 2094.273 < 2.2e-16 \*\*\*

pcb118 1 85429 85429 2097.405 < 2.2e-16 \*\*\*

pcb138 1 62693 62693 1539.202 < 2.2e-16 \*\*\*

pcb180 1 3822 3822 93.834 3.64e-14 \*\*\*

Residuals 64 2607 41

---

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

c.)

> qqnorm(residuals(lm1))

> plot(lm1, which=2)

**11.44**

a.) > plot(lm1)

b.)

> library(readxl)

> pcb2 <- read\_excel("~/Downloads/pcb.xls",

+ sheet = "Sheet1")

> View(pcb2)

> lm1<-lm(pcb~(pcb52+pcb118+pcb138+pcb180), data=pcb2)

> summary(lm1)

Call:

lm(formula = pcb ~ (pcb52 + pcb118 + pcb138 + pcb180), data = pcb2)

Residuals:

Min 1Q Median 3Q Max

-12.2421 -2.1762 -0.1378 1.7036 14.2051

Coefficients:

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

(Intercept) 1.6277 0.8858 1.838 0.0709 .

pcb52 14.4420 0.6960 20.751 < 2e-16 \*\*\*

pcb118 2.5996 0.5164 5.034 4.40e-06 \*\*\*

pcb138 4.0541 0.3752 10.805 6.89e-16 \*\*\*

pcb180 4.1086 0.3175 12.942 < 2e-16 \*\*\*

---

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

Residual standard error: 4.555 on 62 degrees of freedom

Multiple R-squared: 0.9941, Adjusted R-squared: 0.9938

F-statistic: 2629 on 4 and 62 DF, p-value: < 2.2e-16

>plot(lm1)

**11.45**

a.)

> lm2<-lm(pcb~(pcb52+pcb118+pcb138), data=pcb)

> summary(lm2)

Call:

lm(formula = pcb ~ (pcb52 + pcb118 + pcb138), data = pcb)

Residuals:

Min 1Q Median 3Q Max

-29.6219 -3.3502 0.8791 3.3785 29.5217

Coefficients:

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

(Intercept) -1.0184 1.8895 -0.539 0.592

pcb52 12.6442 1.1291 11.198 <2e-16 \*\*\*

pcb118 0.3131 0.8333 0.376 0.708

pcb138 8.2546 0.3279 25.177 <2e-16 \*\*\*

---

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

Residual standard error: 9.945 on 65 degrees of freedom

Multiple R-squared: 0.9732, Adjusted R-squared: 0.972

F-statistic: 786.7 on 3 and 65 DF, p-value: < 2.2e-16

**11.46**

a.)

> lmteq<-lm(teq~(teqpcb+teqdioxin+teqfuran), data=pcb)

> summary(lmteq)

Call:

lm(formula = teq ~ (teqpcb + teqdioxin + teqfuran), data = pcb)

Residuals:

Min 1Q Median 3Q Max

-5.638e-06 -2.844e-06 -1.680e-06 -1.130e-06 3.714e-05

Coefficients:

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

(Intercept) 3.426e-07 1.917e-06 1.790e-01 0.859

teqpcb 1.000e+00 8.239e-07 1.214e+06 <2e-16 \*\*\*

teqdioxin 1.000e+00 1.761e-06 5.677e+05 <2e-16 \*\*\*

teqfuran 1.000e+00 5.664e-06 1.766e+05 <2e-16 \*\*\*

---

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

Residual standard error: 7.95e-06 on 65 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: 1

F-statistic: 9.581e+11 on 3 and 65 DF, p-value: < 2.2e-16

> anova(lmteq)

Analysis of Variance Table

Response: teq

Df Sum Sq Mean Sq F value Pr(>F)

teqpcb 1 152.801 152.801 2.4174e+12 < 2.2e-16 \*\*\*

teqdioxin 1 26.903 26.903 4.2562e+11 < 2.2e-16 \*\*\*

teqfuran 1 1.970 1.970 3.1174e+10 < 2.2e-16 \*\*\*

Residuals 65 0.000 0.000

---

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

**11.47**

> lmteq2<-lm(teq~(pcb52+pcb118+pcb138+pcb180), data=pcb)

> summary(lmteq2)

Call:

lm(formula = teq ~ (pcb52 + pcb118 + pcb138 + pcb180), data = pcb)

Residuals:

Min 1Q Median 3Q Max

-1.6655 -0.6000 -0.1814 0.5162 2.7025

Coefficients:

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

(Intercept) 1.059965 0.184450 5.747 2.73e-07 \*\*\*

pcb52 -0.097277 0.109383 -0.889 0.37716

pcb118 0.306184 0.096388 3.177 0.00229 \*\*

pcb138 0.105786 0.074697 1.416 0.16156

pcb180 -0.003905 0.064784 -0.060 0.95212

---

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

Residual standard error: 0.9576 on 64 degrees of freedom

Multiple R-squared: 0.6769, Adjusted R-squared: 0.6568

F-statistic: 33.53 on 4 and 64 DF, p-value: 4.489e-15

> summary.aov(lmteq2)

Df Sum Sq Mean Sq F value Pr(>F)

pcb52 1 29.85 29.85 32.553 3.21e-07 \*\*\*

pcb118 1 83.61 83.61 91.174 6.30e-14 \*\*\*

pcb138 1 9.52 9.52 10.378 0.00201 \*\*

pcb180 1 0.00 0.00 0.004 0.95212

Residuals 64 58.69 0.92

---

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

> plot(lmteq2)

**11.48**

> library(readxl)

> log <- read\_excel("~/Downloads/pcb.xls",

+ sheet = "log")

> View(log)

> logs<-subset(log, select=c("log138", "log153", "log180", "log28", "log52", "log126", "log118","logpcb","logteq"))

> View(logs)

> summary(logs)

**log138** **log153** **log180**  **log28** **log52**

Min. :-0.1938 Min. :-0.0655 Min. :-0.40340 Min. :-2.2218 Min. :-1.69897

1st Qu.: 0.5024 1st Qu.: 0.5185 1st Qu.: 0.09342 1st Qu.:-0.8996 1st Qu.:-0.64207

Median : 0.6920 Median : 0.7356 Median : 0.42975 Median :-0.5817 Median :-0.32148

Mean : 0.7009 Mean : 0.7397 Mean : 0.42354 Mean :-0.5793 Mean :-0.33537

3rd Qu.: 0.9370 3rd Qu.: 0.9943 3rd Qu.: 0.65225 3rd Qu.:-0.3645 3rd Qu.:-0.04964

Max. : 1.5092 Max. : 1.6385 Max. : 1.49831 Max. : 0.8407 Max. : 0.95713

**log126** **log118** **logpcb** **logteq**

Min. :-2.585 Min. :-0.6271 Min. :0.7853 Min. :-0.02761

1st Qu.:-2.268 1st Qu.: 0.1732 1st Qu.:1.4798 1st Qu.: 0.13274

Median :-2.131 Median : 0.3838 Median :1.6809 Median : 0.31534

Mean :-2.104 Mean : 0.3717 Mean :1.7011 Mean : 0.34950

3rd Qu.:-1.833 3rd Qu.: 0.5899 3rd Qu.:1.9620 3rd Qu.: 0.57174

Max. :-1.499 Max. : 1.2765 Max. :2.5034 Max. : 0.81245

> boxplot(logs)

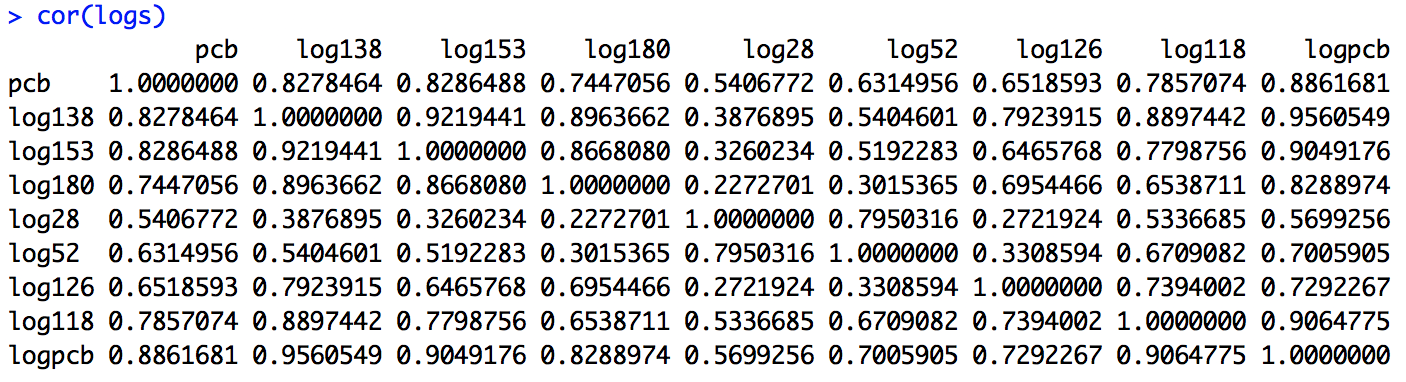
**11.49**

a.)

> logs<-subset(log, select=c("pcb", "log138", "log153", "log180", "log28", "log52", "log126", "log118", "logpcb"))

> pairs(logs, pch=".")

> cor(logs)

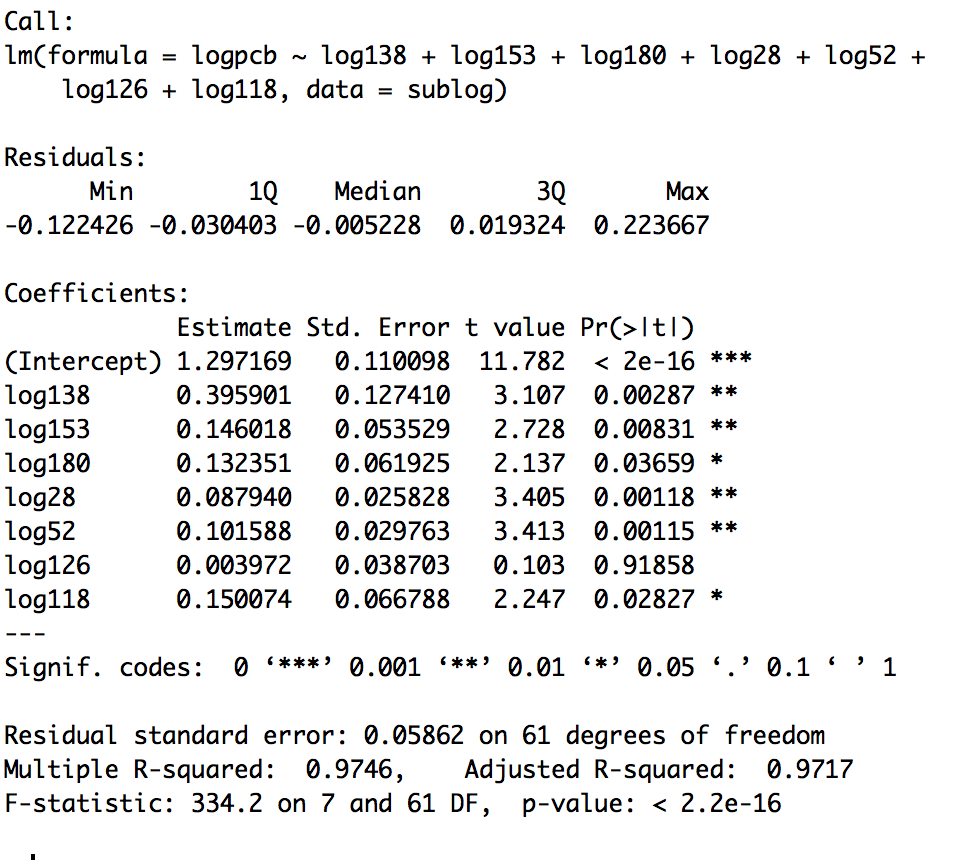


**11.50**

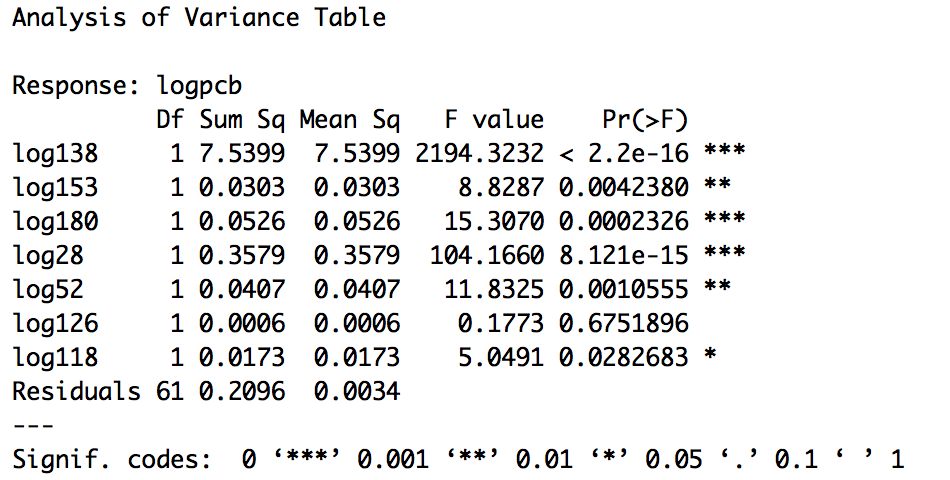
> sublog=subset(log, select=c("logpcb", "log138", "log153", "log180", "log28", "log52", "log126", "log118”))

> lmlog<-lm(logpcb~log138+log153+log180+log28+log52+log126+log118,data=sublog)

> summary(lmlog)



>anova(lmlog)



**11.51**

> sublog=subset(log, select=c("logteq", "logpcb", "log138", "log153", "log180", "log28", "log52", "log126", "log118"))

> lmlog<-lm(logteq~log138+log153+log180+log28+log52+log126+log118+logpcb,data=sublog)

> summary(lmlog)

