Question 1

**> library(faraway)**

**> data(seatpos)**

(a)

**> model1.1<-lm(hipcenter~Age+Weight+HtShoes+Ht+Seated+Arm+Thigh+Leg,data=seatpos)**

**> summary(model1.1)**

Call:

lm(formula = hipcenter ~ Age + Weight + HtShoes + Ht + Seated +

Arm + Thigh + Leg, data = seatpos)

Residuals:

Min 1Q Median 3Q Max

-73.827 -22.833 -3.678 25.017 62.337

Coefficients:

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

(Intercept) 436.43213 166.57162 2.620 0.0138 \*

Age 0.77572 0.57033 1.360 0.1843

Weight 0.02631 0.33097 0.080 0.9372

HtShoes -2.69241 9.75304 -0.276 0.7845

Ht 0.60134 10.12987 0.059 0.9531

Seated 0.53375 3.76189 0.142 0.8882

Arm -1.32807 3.90020 -0.341 0.7359

Thigh -1.14312 2.66002 -0.430 0.6706

Leg -6.43905 4.71386 -1.366 0.1824

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 37.72 on 29 degrees of freedom

Multiple R-squared: 0.6866, Adjusted R-squared: 0.6001

F-statistic: 7.94 on 8 and 29 DF, p-value: 1.306e-05

(b)

None of the variables appears to be significant based on the individual t-test for coefficient. But since we have a very small p-value for the overall F-test shows that at least one of the variables is significant.

(c)

**>vif(model1)**

Age Weight HtShoes Ht Seated Arm Thigh

1.997931 3.647030 307.429378 333.137832 8.951054 4.496368 2.762886

Leg

6.694291

(d)

**>model1.2<-lm(hipcenter~Age+Weight+Seated+Arm+Thigh+Leg,data=seatpos)**

**>summary(model2)**

Call:

lm(formula = hipcenter ~ Age + Weight + Seated + Arm + Thigh +

Leg, data = seatpos)

Residuals:

Min 1Q Median 3Q Max

-68.296 -23.340 -5.672 24.183 74.065

Coefficients:

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

(Intercept) 409.00851 159.49517 2.564 0.0154 \*

Age 0.83110 0.52771 1.575 0.1254

Weight -0.03251 0.31254 -0.104 0.9178

Seated -1.73576 2.48225 -0.699 0.4896

Arm -2.00541 3.69731 -0.542 0.5914

Thigh -1.91970 2.24858 -0.854 0.3998

Leg -8.40876 3.91939 -2.145 0.0399 \*

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 36.91 on 31 degrees of freedom

Multiple R-squared: 0.6791, Adjusted R-squared: 0.617

F-statistic: 10.94 on 6 and 31 DF, p-value: 1.571e-06

(e)

Except variable leg, other variables all appear to be not significant based on the individual t-test. And the p-value for overall F-test is very small, which suggests that at least one of the variables is significant.

(f)

**> vif(model1.2)**

Age Weight Seated Arm Thigh Leg

1.786192 3.396124 4.069626 4.219519 2.061632 4.832701

Question 2

**>Age<-c(1,2,3,4,5)**

**>Number<-c(123,78,32,17,24)**

**>Average\_Weight<-c(7.9725,7.9503,7.9276,7.8962,7.8730)**

**>sd<-c(0.01409,0.02272,0.03426,0.04057,0.05353)**

(a) model2.1 is unweighted regression

**>model2.1<-lm(Average\_Weight~Age)**

**>summary(model2.1)**

Call:

lm(formula = Average\_Weight ~ Age)

Residuals:

1 2 3 4 5

-0.00204 0.00107 0.00368 -0.00241 -0.00030

Coefficients:

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

(Intercept) 7.9998500 0.0030123 2655.74 1.18e-10 \*\*\*

Age -0.0253100 0.0009082 -27.87 0.000101 \*\*\*

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 0.002872 on 3 degrees of freedom

Multiple R-squared: 0.9962, Adjusted R-squared: 0.9949

F-statistic: 776.6 on 1 and 3 DF, p-value: 0.0001014

(b)

**>confint(model2.1)**

2.5 % 97.5 %

(Intercept) 7.99026354 8.00943646

Age -0.02820043 -0.02241957

(c) model2.2 is weighted regression

**>model2.2<-lm(Average\_Weight~Age,weights=Number)**

**>summary(model2.2)**

Call:

lm(formula = Average\_Weight ~ Age, weights = Number)

Weighted Residuals:

1 2 3 4 5

-0.011442 0.012903 0.019537 -0.013415 -0.008627

Coefficients:

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

(Intercept) 7.9982244 0.0020332 3933.9 3.62e-11 \*\*\*

Age -0.0246927 0.0008427 -29.3 8.73e-05 \*\*\*

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 0.01764 on 3 degrees of freedom

Multiple R-squared: 0.9965, Adjusted R-squared: 0.9954

F-statistic: 858.6 on 1 and 3 DF, p-value: 8.729e-05

(d)

**>confint(model2.2)**

2.5 % 97.5 %

(Intercept) 7.9917539 8.00469482

Age -0.0273745 -0.02201086

(e)

**> model2.3<-lm(Average\_Weight~Age,weights=(Number/sd^2))**

**> summary(model2.3)**

Call:

lm(formula = Average\_Weight ~ Age, weights = (Number/sd^2))

Weighted Residuals:

1 2 3 4 5

-0.2091 0.5017 0.3875 -0.5383 -0.4339

Coefficients:

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

(Intercept) 7.9965218 0.0013220 6049 9.96e-12 \*\*\*

Age -0.0237562 0.0008797 -27 0.000111 \*\*\*

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 0.5549 on 3 degrees of freedom

Multiple R-squared: 0.9959, Adjusted R-squared: 0.9945

F-statistic: 729.2 on 1 and 3 DF, p-value: 0.0001114

(f)

**> confint(model2.3)**

2.5 % 97.5 %

(Intercept) 7.99231466 8.00072893

Age -0.02655593 -0.02095642

Question 3

**>data(salmonella)**

(a)

**>model3<-lm(colonies~dose,data=salmonella)**

**>summary(model3)**

Call:

lm(formula = colonies ~ dose, data = salmonella)

Residuals:

Min 1Q Median 3Q Max

-13.488 -8.991 -1.569 7.569 31.736

Coefficients:

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

(Intercept) 27.683154 3.462169 7.996 5.58e-07 \*\*\*

dose 0.005805 0.008006 0.725 0.479

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

Residual standard error: 12.08 on 16 degrees of freedom

Multiple R-squared: 0.03181, Adjusted R-squared: -0.0287

F-statistic: 0.5257 on 1 and 16 DF, p-value: 0.4789

(b) Because the we have independent replications at each of the doses

(c)

**>linmod<-lm(colonies~dose,data=salmonella)**

**>genmod<-lm(colonies~factor(dose),data=salmonella)**

**>anova(linmod,genmod)**

Analysis of Variance Table

Model 1: colonies ~ dose

Model 2: colonies ~ factor(dose)

Res.Df RSS Df Sum of Sq F Pr(>F)

1 16 2335.1

2 12 1091.3 4 1243.7 3.4189 0.04373 \*

---

Signif. codes: 0 ¡®\*\*\*¡¯ 0.001 ¡®\*\*¡¯ 0.01 ¡®\*¡¯ 0.05 ¡®.¡¯ 0.1 ¡® ¡¯ 1

(d)

Since the F is 3.4189, which is very small, alongwith a less than 0.05 p-value, we fail to reject the null hypothesis. Therefore, the mean model is correct.

(e)

**>pure\_error\_estimate = 1091.3/4**

**>pure\_error\_estimate**

[1] 272.825