Homework 5

Devin Etcitty

STAT4205

dce2108

```
library(alr4)
```

Problem 6.1

UN11 dataframe

```
model6.6 <- lm(lifeExpF ~ 1, data=UN11)
model6.7 <- lm(lifeExpF ~ group,  data=UN11)



summary(model6.6)
```

```
##
## Call:
## lm(formula = lifeExpF ~ 1, data = UN11)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.183  -6.633   3.597   7.292  14.827
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  72.2932     0.7177   100.7   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.12 on 198 degrees of freedom
```

```
summary(model6.7)
```

```
##
## Call:
## lm(formula = lifeExpF ~ group, data = UN11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.8367  -3.3045   0.3635   2.7183  18.2277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   82.446      1.128  73.095  < 2e-16 ***
## groupother    -7.120      1.271  -5.602  7.1e-08 ***
## groupafrica  -22.674      1.420 -15.968  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.28 on 196 degrees of freedom
## Multiple R-squared:  0.6191, Adjusted R-squared:  0.6152
## F-statistic: 159.3 on 2 and 196 DF,  p-value: < 2.2e-16
```

```
anova(model6.6, model6.7)
```

```
## Analysis of Variance Table
##
## Model 1: lifeExpF ~ 1
## Model 2: lifeExpF ~ group
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1    198 20293.2
## 2    196  7730.2  2     12563 159.27 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Problem 6.2

```
model6.7 <- lm(lifeExpF ~ group,  data=UN11)
model6.8 <- lm(lifeExpF ~ log(ppgdp),  data=UN11)

summary(model6.7)
```

```
##
## Call:
## lm(formula = lifeExpF ~ group, data = UN11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.8367  -3.3045   0.3635   2.7183  18.2277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   82.446      1.128  73.095  < 2e-16 ***
## groupother    -7.120      1.271  -5.602  7.1e-08 ***
## groupafrica  -22.674      1.420 -15.968  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.28 on 196 degrees of freedom
## Multiple R-squared:  0.6191, Adjusted R-squared:  0.6152
## F-statistic: 159.3 on 2 and 196 DF,  p-value: < 2.2e-16
```

```
summary(model6.8)
```

```
##
## Call:
## lm(formula = lifeExpF ~ log(ppgdp), data = UN11)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -25.749  -2.879   1.280   3.987  12.345
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  29.8148     2.5314   11.78   <2e-16 ***
## log(ppgdp)    5.0188     0.2942   17.06   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.448 on 197 degrees of freedom
## Multiple R-squared:  0.5964, Adjusted R-squared:  0.5943
## F-statistic: 291.1 on 1 and 197 DF,  p-value: < 2.2e-16
```

```
anova(model6.7, model6.8)
```

```
## Analysis of Variance Table
##
## Model 1: lifeExpF ~ group
## Model 2: lifeExpF ~ log(ppgdp)
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    196 7730.2
## 2    197 8190.7 -1   -460.49 11.676 0.0007699 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Problem 6.3

```
model6.7 <- lm(lifeExpF ~ group,  data=UN11)
model6.9 <- lm(lifeExpF ~ group + log(ppgdp),  data=UN11)

summary(model6.7)
```

```
##
## Call:
## lm(formula = lifeExpF ~ group, data = UN11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.8367  -3.3045   0.3635   2.7183  18.2277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   82.446      1.128  73.095  < 2e-16 ***
## groupother    -7.120      1.271  -5.602  7.1e-08 ***
## groupafrica  -22.674      1.420 -15.968  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.28 on 196 degrees of freedom
## Multiple R-squared:  0.6191, Adjusted R-squared:  0.6152
## F-statistic: 159.3 on 2 and 196 DF,  p-value: < 2.2e-16
```

```
summary(model6.9)
```

```
##
## Call:
## lm(formula = lifeExpF ~ group + log(ppgdp), data = UN11)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -18.6348  -2.1741   0.2441   2.3537  14.6539
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   49.529      3.400  14.569  < 2e-16 ***
## groupother    -1.535      1.174  -1.308    0.193
## groupafrica  -12.170      1.557  -7.814 3.35e-13 ***
## log(ppgdp)     3.177      0.316  10.056  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.109 on 195 degrees of freedom
## Multiple R-squared:  0.7492, Adjusted R-squared:  0.7453
## F-statistic: 194.1 on 3 and 195 DF,  p-value: < 2.2e-16
```

```
anova(model6.7, model6.9)
```

```
## Analysis of Variance Table
##
## Model 1: lifeExpF ~ group
## Model 2: lifeExpF ~ group + log(ppgdp)
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    196 7730.2
## 2    195 5090.4  1    2639.8 101.12 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(model6.9, model6.7)
```

```
## Analysis of Variance Table
##
## Model 1: lifeExpF ~ group + log(ppgdp)
## Model 2: lifeExpF ~ group
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    195 5090.4
## 2    196 7730.2 -1   -2639.8 101.12 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Problem 6.4

UN11

```
nullHyp <- lm(lifeExpF ~ log(ppgdp) + group:log(ppgdp), data=UN11)
alterHyp <- lm(lifeExpF ~ group + log(ppgdp) + group:log(ppgdp), data=UN11)

summary(nullHyp)
```

```
##
## Call:
## lm(formula = lifeExpF ~ log(ppgdp) + group:log(ppgdp), data = UN11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.6121  -2.5029   0.3037   2.4489  15.3486
```

```
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              43.8040     2.6231  16.699  < 2e-16 ***
## log(ppgdp)                3.7245     0.2677  13.912  < 2e-16 ***
## log(ppgdp):groupother    -0.0698     0.1153  -0.605    0.546
## log(ppgdp):groupafrica   -1.4303     0.1726  -8.285 1.87e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.18 on 195 degrees of freedom
## Multiple R-squared:  0.7422, Adjusted R-squared:  0.7382
## F-statistic: 187.1 on 3 and 195 DF,  p-value: < 2.2e-16
```

```
summary(alterHyp)
```

```
##
## Call:
## lm(formula = lifeExpF ~ group + log(ppgdp) + group:log(ppgdp),
##     data = UN11)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -18.634  -2.089   0.301   2.255  14.489
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              59.2137    15.2203   3.890 0.000138 ***
## groupother              -11.1731    15.5948  -0.716 0.474572
## groupafrica             -22.9848    15.7838  -1.456 0.146954
## log(ppgdp)                2.2425     1.4664   1.529 0.127844
## groupother:log(ppgdp)     0.9294     1.5177   0.612 0.540986
## groupafrica:log(ppgdp)    1.0950     1.5785   0.694 0.488703
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.129 on 193 degrees of freedom
## Multiple R-squared:  0.7498, Adjusted R-squared:  0.7433
## F-statistic: 115.7 on 5 and 193 DF,  p-value: < 2.2e-16
```

Problem 6.5

UN11

```
paraReg <- lm(lifeExpF ~ group + log(ppgdp), data = UN11)
summary(paraReg)$coeff
```

```
##                Estimate Std. Error   t value     Pr(>|t|)
## (Intercept)  49.529241  3.3995539 14.569336 5.142392e-33
## groupother   -1.534683  1.1736824 -1.307579 1.925556e-01
## groupafrica -12.170365  1.5574486 -7.814297 3.352679e-13
## log(ppgdp)    3.177320  0.3159597 10.056092 1.972779e-19
```

```
summary(paraReg)
```

```
##
```

```
## Call:
## lm(formula = lifeExpF ~ group + log(ppgdp), data = UN11)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -18.6348  -2.1741   0.2441   2.3537  14.6539
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   49.529      3.400  14.569  < 2e-16 ***
## groupother    -1.535      1.174  -1.308    0.193
## groupafrica  -12.170      1.557  -7.814 3.35e-13 ***
## log(ppgdp)     3.177      0.316  10.056  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.109 on 195 degrees of freedom
## Multiple R-squared:  0.7492, Adjusted R-squared:  0.7453
## F-statistic: 194.1 on 3 and 195 DF,  p-value: < 2.2e-16
```

Problem 6.7

fuel2001

```
model6.22 <- lm(FuelC ~ Tax + Drivers + Income + log(Miles), data=fuel2001)
model6.23 <- lm(FuelC~ log(Miles) + Income + Drivers + Tax, data=fuel2001)

summary(model6.22)
```

```
##
## Call:
## lm(formula = FuelC ~ Tax + Drivers + Income + log(Miles), data = fuel2001)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1676904  -126002   -21638   146118  1849371
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.274e+05  1.219e+06   0.186    0.853
## Tax         -2.270e+04  1.436e+04  -1.581    0.121
## Drivers      6.566e-01  2.198e-02  29.868   <2e-16 ***
## Income      -1.820e+01  1.745e+01  -1.043    0.302
## log(Miles)   7.579e+04  8.503e+04   0.891    0.377
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 453600 on 46 degrees of freedom
## Multiple R-squared:  0.974,  Adjusted R-squared:  0.9717
## F-statistic: 430.6 on 4 and 46 DF,  p-value: < 2.2e-16
```

```
summary(model6.23)
```

```
##
## Call:
```

```
## lm(formula = FuelC ~ log(Miles) + Income + Drivers + Tax, data = fuel2001)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1676904  -126002   -21638   146118  1849371
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.274e+05  1.219e+06   0.186    0.853
## log(Miles)   7.579e+04  8.503e+04   0.891    0.377
## Income      -1.820e+01  1.745e+01  -1.043    0.302
## Drivers      6.566e-01  2.198e-02  29.868   <2e-16 ***
## Tax         -2.270e+04  1.436e+04  -1.581    0.121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 453600 on 46 degrees of freedom
## Multiple R-squared:  0.974,  Adjusted R-squared:  0.9717
## F-statistic: 430.6 on 4 and 46 DF,  p-value: < 2.2e-16
```

```
anova(model6.22, model6.23)
```

```
## Analysis of Variance Table
##
## Model 1: FuelC ~ Tax + Drivers + Income + log(Miles)
## Model 2: FuelC ~ log(Miles) + Income + Drivers + Tax
##   Res.Df       RSS Df Sum of Sq F Pr(>F)
## 1     46 9.463e+12
## 2     46 9.463e+12  0 0.0039062
```

Problem 6.14

MinnLand

```
model <- lm(log(acrePrice) ~ year, data=MinnLand)
summary(model)
```

```
##
## Call:
## lm(formula = log(acrePrice) ~ year, data = MinnLand)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1008 -0.3773  0.1285  0.4365  2.2624
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.939e+02  3.984e+00  -48.67   <2e-16 ***
## year         1.005e-01  1.985e-03   50.60   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6808 on 18698 degrees of freedom
## Multiple R-squared:  0.1204, Adjusted R-squared:  0.1204
## F-statistic:  2560 on 1 and 18698 DF,  p-value: < 2.2e-16
```

Problem 6.18

wm2

```
plot(wm2)
```