

Final Project Appendix

Parker Johnson, Teagan Johnson

11/23/2021

Appendix:

The following is the R code used to reach our conclusions.

First, we checked which years would be appropriate to use in our model. The high number of missing values for 2015 indicated that this year would not be the most accurate fit, and checking the number of missing values and inaccurate 0 values for 2014, we found that this year would be much better to use.

```
LifeExpectancy <- read.csv("Life_Expectancy_Data.csv")

vec2015 <- c(2015)

LifeExpectancy2015Bad <- LifeExpectancy[LifeExpectancy$Year %in% vec2015, ]
sapply(LifeExpectancy2015Bad, function(x) sum(is.na(x)))
```

```
##          Country          Year
##          0          0
##          Status      Life.expectancy
##          0          0
##      Adult.Mortality      infant.deaths
##          0          0
##          Alcohol      percentage.expenditure
##          177          0
##      Hepatitis.B          Measles
##          9          0
##          BMI      under.five.deaths
##          2          0
##          Polio      Total.expenditure
##          0          181
##      Diphtheria      HIV.AIDS
##          0          0
##          GDP      Population
##          29          41
##      thinness..1.19.years      thinness.5.9.years
##          2          2
##      Income.composition.of.resources      Schooling
```

```
##                                10                                10

vec <- c(2014)
LifeExpectancy2014Bad <- LifeExpectancy[LifeExpectancy$Year %in% vec, ]
sapply(LifeExpectancy2014Bad, function(x) sum(is.na(x)))
```

```
##                Country                Year
##                0                0
##                Status                Life.expectancy
##                0                0
##                Adult.Mortality                infant.deaths
##                0                0
##                Alcohol                percentage.expenditure
##                1                0
##                Hepatitis.B                Measles
##                10                0
##                BMI                under.five.deaths
##                2                0
##                Polio                Total.expenditure
##                0                2
##                Diphtheria                HIV.AIDS
##                0                0
##                GDP                Population
##                28                41
##                thinness..1.19.years                thinness.5.9.years
##                2                2
## Income.composition.of.resources                Schooling
##                10                10
```

```
LifeExpectancy2014Old <- na.omit(LifeExpectancy[LifeExpectancy$Year %in% vec, ])
sapply(LifeExpectancy2014Old, function(x) sum(x == 0))
```

```
##                Country                Year
##                0                0
##                Status                Life.expectancy
##                0                0
##                Adult.Mortality                infant.deaths
##                0                34
##                Alcohol                percentage.expenditure
##                0                0
##                Hepatitis.B                Measles
##                0                53
##                BMI                under.five.deaths
##                0                30
##                Polio                Total.expenditure
```

```
##              0              0
##          Diphtheria          HIV.AIDS
##              0              0
##              GDP              Population
##              0              0
##      thinness..1.19.years      thinness.5.9.years
##              0              0
## Income.composition.of.resources      Schooling
##              0              0
```

```
LifeExpectancy2014 <- subset(LifeExpectancy2014Old, select =
                             -c(infant.deaths, Measles, under.five.deaths))
```

Since infant deaths, measles, and under 5 deaths contained a high number of unreasonable 0 values, we decided that it was better to remove these terms altogether, since removing the countries that had these 0 values led to models that were less accurate and contained more variation. Our starting model for 2014 is listed below:

```
LifeExpectancy2014_lm <- lm(Life.expectancy ~ Status + Adult.Mortality +
                             Alcohol + percentage.expenditure + BMI +
                             Hepatitis.B + Polio + Total.expenditure +
                             Diphtheria + HIV.AIDS + GDP + Population +
                             thinness..1.19.years + thinness.5.9.years +
                             Income.composition.of.resources +
                             Schooling, data = LifeExpectancy2014)
summary(LifeExpectancy2014_lm)
```

```
##
## Call:
## lm(formula = Life.expectancy ~ Status + Adult.Mortality + Alcohol +
##     percentage.expenditure + BMI + Hepatitis.B + Polio + Total.expenditure +
##     Diphtheria + HIV.AIDS + GDP + Population + thinness..1.19.years +
##     thinness.5.9.years + Income.composition.of.resources + Schooling,
##     data = LifeExpectancy2014)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.164  -1.866   0.114   1.869   8.347
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.034e+01  3.270e+00  15.395  < 2e-16 ***
## StatusDeveloping  -1.104e+00  1.027e+00  -1.075  0.284578
## Adult.Mortality   -1.778e-02  4.133e-03  -4.300  3.61e-05 ***
## Alcohol           -4.175e-03  9.656e-02  -0.043  0.965589
```

```
## percentage.expenditure      4.805e-04  4.602e-04   1.044 0.298688
## BMI                         -6.422e-03  1.976e-02  -0.325 0.745798
## Hepatitis.B                 1.003e-02  2.797e-02   0.359 0.720515
## Polio                       -9.790e-03  2.114e-02  -0.463 0.644083
## Total.expenditure           2.827e-01  1.273e-01   2.220 0.028399 *
## Diphtheria                  1.226e-02  3.424e-02   0.358 0.720824
## HIV.AIDS                    -8.753e-01  2.450e-01  -3.573 0.000519 ***
## GDP                         -6.382e-05  6.622e-05  -0.964 0.337210
## Population                  1.342e-09  2.853e-09   0.470 0.638923
## thinness..1.19.years        5.313e-02  1.907e-01   0.279 0.780989
## thinness.5.9.years          -1.400e-01  1.921e-01  -0.729 0.467504
## Income.composition.of.resources 3.699e+01  6.165e+00   6.000 2.39e-08 ***
## Schooling                    -1.665e-01  2.701e-01  -0.616 0.538897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.186 on 114 degrees of freedom
## Multiple R-squared:  0.8798, Adjusted R-squared:  0.8629
## F-statistic: 52.13 on 16 and 114 DF,  p-value: < 2.2e-16
```

```
vif(LifeExpectancy2014_lm)
```

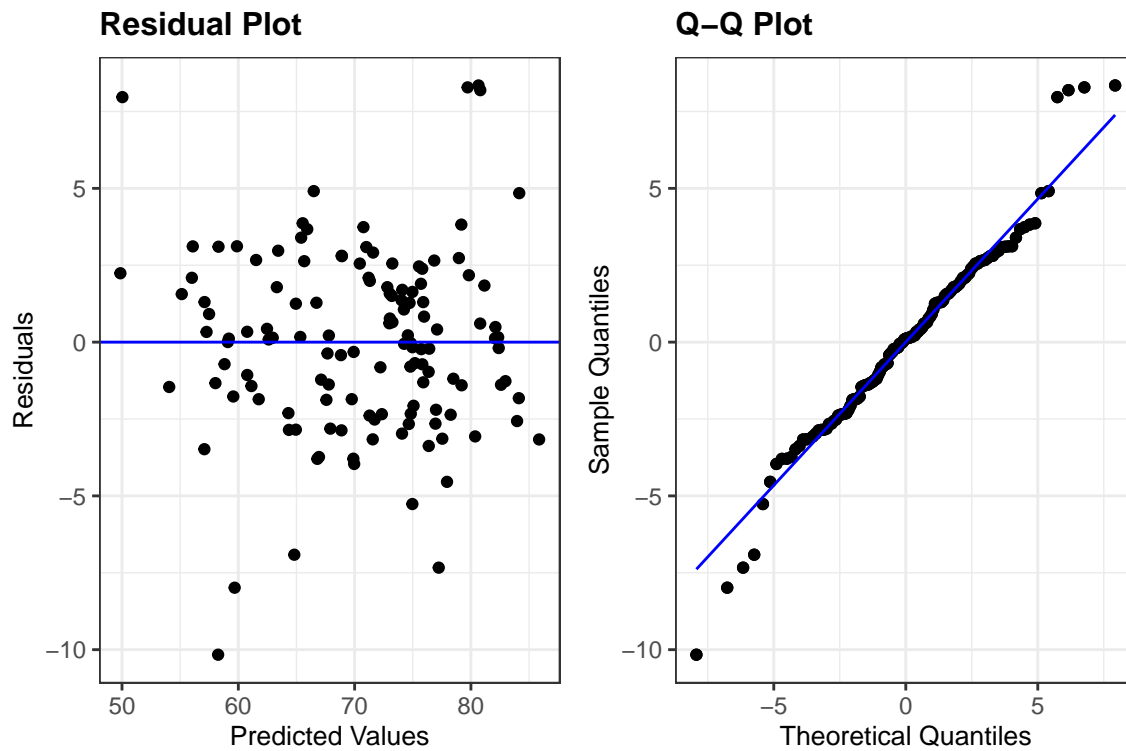
```
##              Status              Adult.Mortality
##      1.687415              2.653753
##      Alcohol      percentage.expenditure
##      1.997310              11.635049
##      BMI              Hepatitis.B
##      2.149280              5.655938
##      Polio      Total.expenditure
##      2.514107              1.332191
##      Diphtheria      HIV.AIDS
##      7.157811              1.874949
##      GDP      Population
##      12.199568              1.419649
##      thinness..1.19.years      thinness.5.9.years
##      9.094802              9.752426
##      Income.composition.of.resources      Schooling
##      11.141152              7.065109
```

Next, we checked residual plots and cook's distance to determine if there were any outliers or necessary transformations present in our data. Since many of the terms currently violate the MLR assumption of constant variance, many transformations are needed.

```
NewLifeExpectancy2014_lm <- lm(Life.expectancy ~ Status + Adult.Mortality +
                               Alcohol + log(percentage.expenditure) + BMI +
```

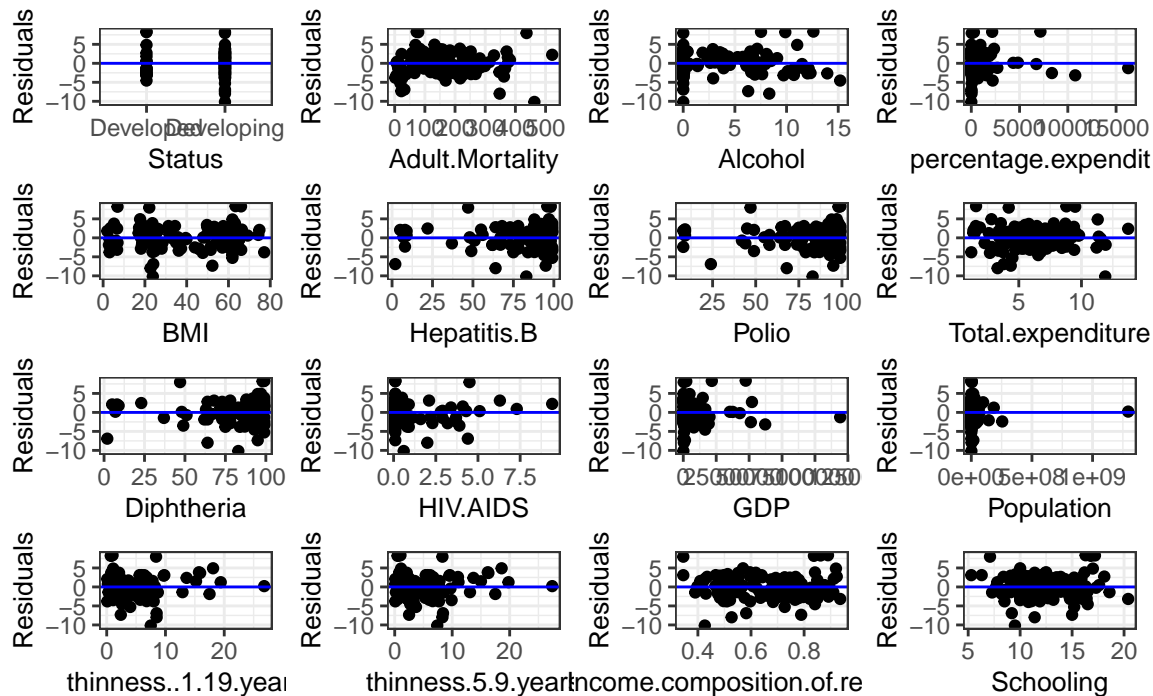
```
Hepatitis.B + Polio + Total.expenditure +  
Diphtheria + HIV.AIDS + GDP + Population +  
thinness..1.19.years + thinness.5.9.years +  
Income.composition.of.resources +  
Schooling, data = LifeExpectancy2014)
```

```
resid_panel(LifeExpectancy2014_lm, plots = c("resid", "qq"))
```

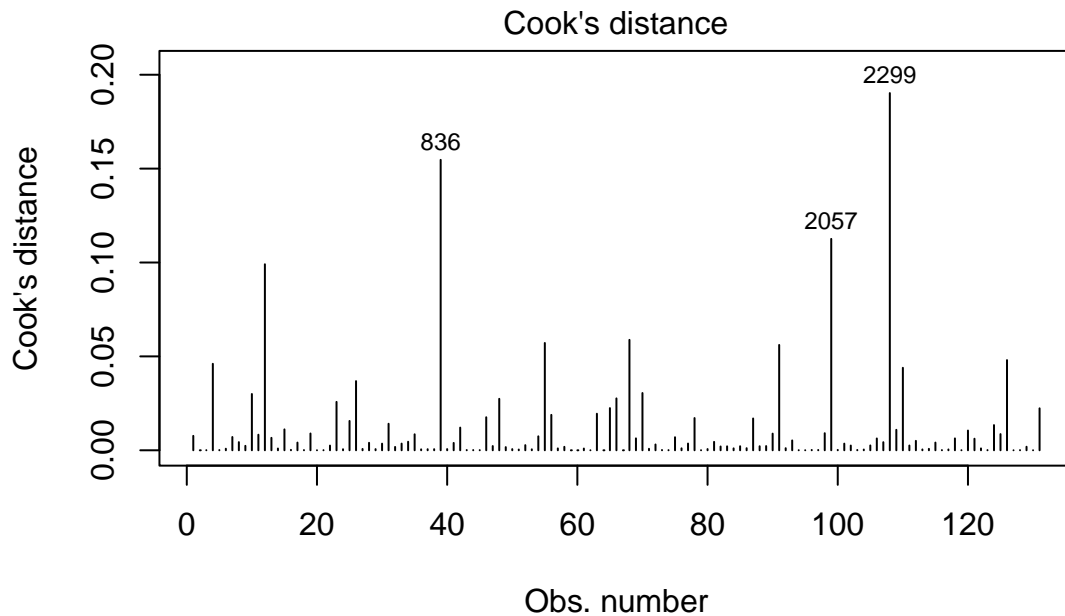


```
resid_xpanel(LifeExpectancy2014_lm)
```

Plots of Residuals vs Predictor Variables



```
plot(NewLifeExpectancy2014_lm, which = 4)
```



$\text{lm}(\text{Life.expectancy} \sim \text{Status} + \text{Adult.Mortality} + \text{Alcohol} + \log(\text{percentage.ex} \dots)$

After transforming many of the variables, the residual plots look much better, and the MLR assumptions are now met. Additionally, since all of the countries had Cook's distances less than 0.2, we concluded that there were not any influential cases that had to be removed.

```
life_exp_2014_lm_mod <- lm(Life.expectancy ~ Status + Adult.Mortality +  
  Alcohol + log(percentage.expenditure) + BMI +
```

```

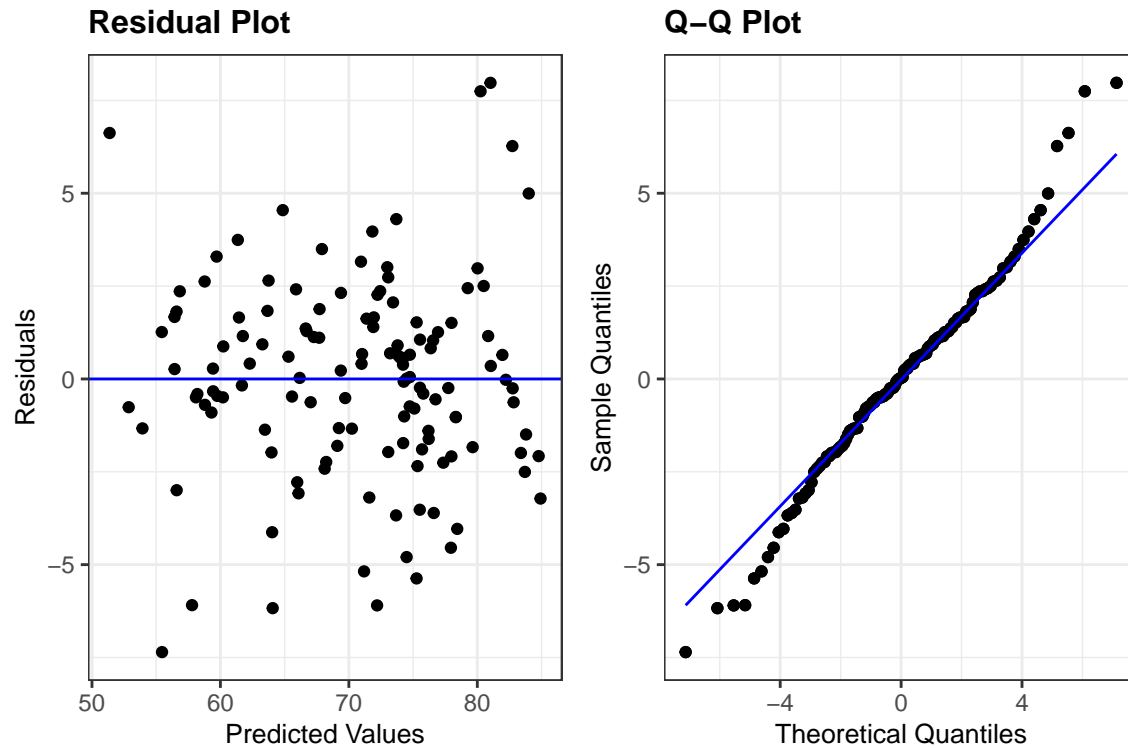
I(Hepatitis.B^2) + I(Polio^2) + Total.expenditure +
I(Diphtheria^2) + log(HIV.AIDS) + log(GDP) +
log(Population) + log(thinness..1.19.years) +
log(thinness.5.9.years) +
Income.composition.of.resources + Schooling,
data = LifeExpectancy2014)

```

```

resid_panel(life_exp_2014_lm_mod, plots = c("resid", "qq"))

```

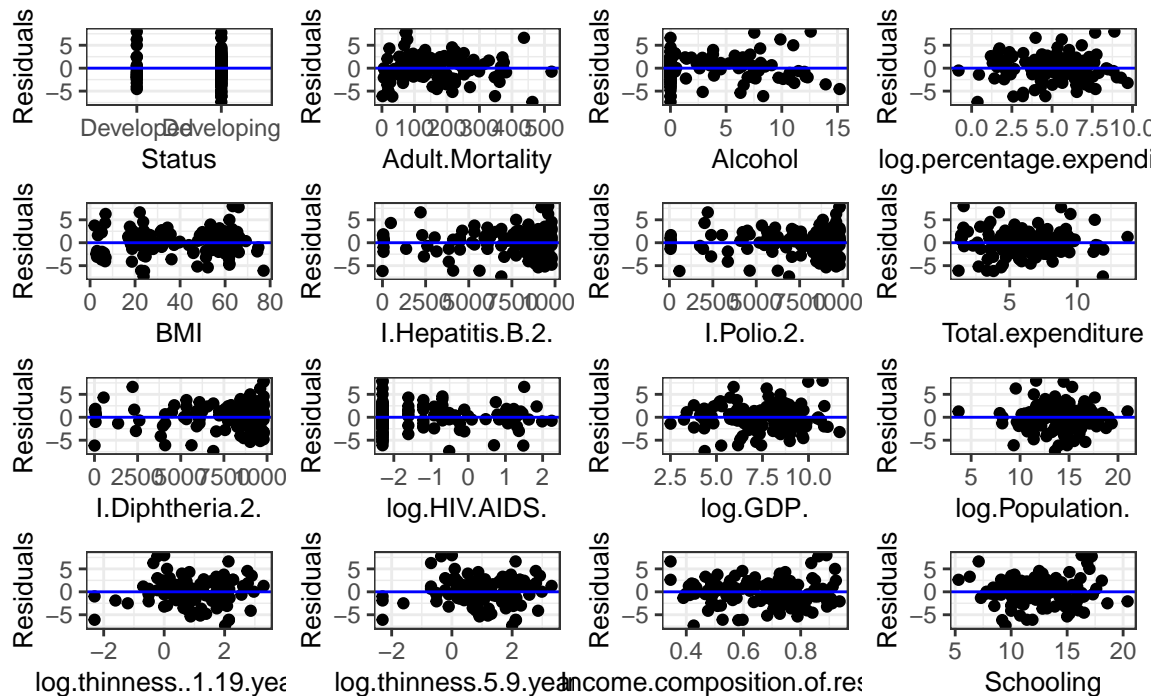


```

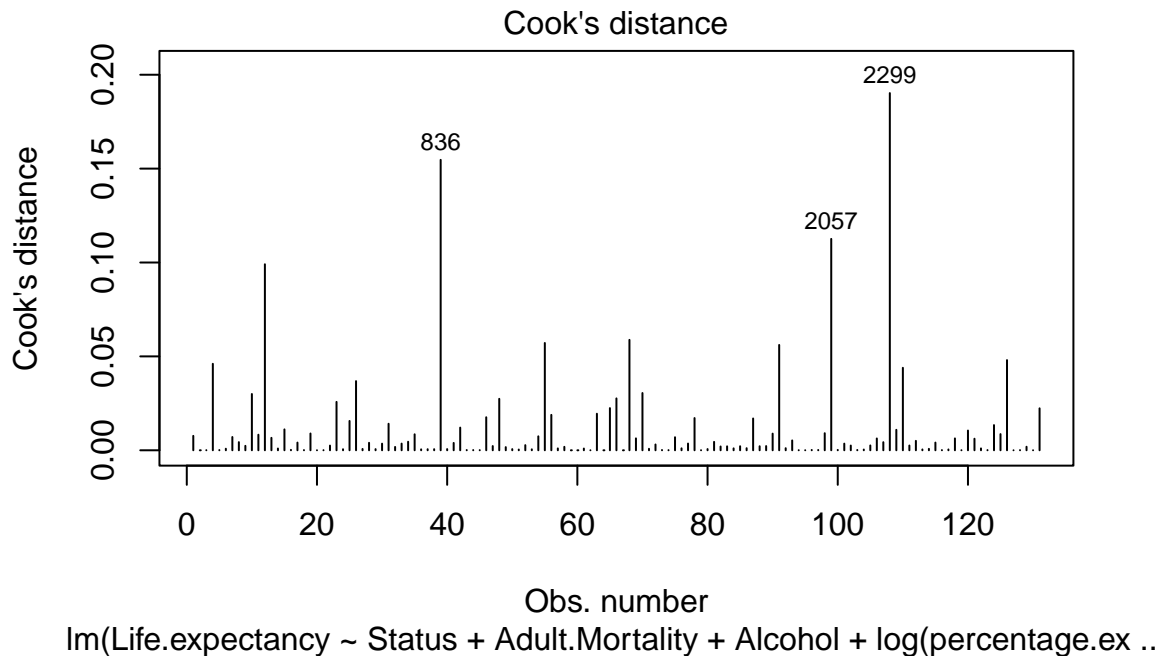
resid_xpanel(life_exp_2014_lm_mod)

```

Plots of Residuals vs Predictor Variables



```
plot(NewLifeExpectancy2014_lm, which = 4)
```



We now have our transformed model for 2014 listed below.

```
life_exp_2014_lm_mod <- lm(Life.expectancy ~ Status + Adult.Mortality +  
  Alcohol + log(percentage.expenditure) + BMI +  
  I(Hepatitis.B^2) + I(Polio^2) + Total.expenditure +  
  I(Diphtheria^2) + log(HIV.AIDS) + log(GDP) +
```



```

log(Population) + log(thinness..1.19.years) +
log(thinness.5.9.years) +
Income.composition.of.resources + Schooling,
data = LifeExpectancy2014)
summary(life_exp_2014_lm_mod)

##
## Call:
## lm(formula = Life.expectancy ~ Status + Adult.Mortality + Alcohol +
##     log(percentage.expenditure) + BMI + I(Hepatitis.B^2) + I(Polio^2) +
##     Total.expenditure + I(Diphtheria^2) + log(HIV.AIDS) + log(GDP) +
##     log(Population) + log(thinness..1.19.years) + log(thinness.5.9.years) +
##     Income.composition.of.resources + Schooling, data = LifeExpectancy2014)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.3548 -1.5550  0.0278  1.5169  7.9731
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.630e+01  3.370e+00  16.706 < 2e-16 ***
## StatusDeveloping -7.777e-01  9.157e-01  -0.849  0.397520
## Adult.Mortality  -1.781e-02  3.676e-03  -4.844  4.05e-06 ***
## Alcohol           4.069e-02  8.633e-02   0.471  0.638300
## log(percentage.expenditure)  1.357e+00  3.917e-01   3.466  0.000747 ***
## BMI              -3.265e-02  1.838e-02  -1.776  0.078373 .
## I(Hepatitis.B^2)   1.128e-04  2.318e-04   0.487  0.627444
## I(Polio^2)         -2.699e-05  1.930e-04  -0.140  0.889053
## Total.expenditure  2.293e-01  1.117e-01   2.053  0.042407 *
## I(Diphtheria^2)    4.074e-05  2.979e-04   0.137  0.891451
## log(HIV.AIDS)      -1.562e+00  3.136e-01  -4.979  2.29e-06 ***
## log(GDP)           -1.279e+00  4.356e-01  -2.936  0.004020 **
## log(Population)    2.472e-02  1.043e-01   0.237  0.813053
## log(thinness..1.19.years)  9.893e-02  7.691e-01   0.129  0.897875
## log(thinness.5.9.years)  -9.409e-01  7.693e-01  -1.223  0.223815
## Income.composition.of.resources  3.171e+01  6.061e+00   5.232  7.71e-07 ***
## Schooling          -2.654e-01  2.495e-01  -1.064  0.289736
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.864 on 114 degrees of freedom
## Multiple R-squared:  0.9029, Adjusted R-squared:  0.8892
## F-statistic: 66.23 on 16 and 114 DF, p-value: < 2.2e-16

```

```
vif(life_exp_2014_lm_mod)
```

```
##              Status              Adult.Mortality
##          1.660790              2.598015
##          Alcohol    log(percentage.expenditure)
##          1.976327              10.357781
##              BMI              I(Hepatitis.B^2)
##          2.302140              6.554729
##          I(Polio^2)              Total.expenditure
##          4.069259              1.269454
##          I(Diphtheria^2)          log(HIV.AIDS)
##          9.570596              2.782027
##          log(GDP)              log(Population)
##          10.536026              1.215480
##          log(thinness..1.19.years)    log(thinness.5.9.years)
##          10.053343              10.951463
## Income.composition.of.resources      Schooling
##          13.331921              7.465800
```

After removing the terms that showed high collinearity or were not significant, our final model is shown below, with all terms being statistically significant with 95% confidence. The p-values for Total.expenditure (0.058) and log(thinness.5.9.years) (0.066) we deemed were close enough to warrant not removing them from our model.

```
life_exp_2014_lm_mod_new <- lm(Life.expectancy ~ Status + Adult.Mortality +
                                log(percentage.expenditure + .5) +
                                Total.expenditure + log(HIV.AIDS) +
                                log(thinness.5.9.years) + Schooling,
                                data = LifeExpectancy2014)
summary(life_exp_2014_lm_mod_new)
```

```
##
## Call:
## lm(formula = Life.expectancy ~ Status + Adult.Mortality + log(percentage.expenditure +
##      0.5) + Total.expenditure + log(HIV.AIDS) + log(thinness.5.9.years) +
##      Schooling, data = LifeExpectancy2014)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.6959 -1.6055  0.2676  1.5830  7.7686
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    59.48671    2.733508   21.762  < 2e-16 ***
## StatusDeveloping    -2.16393    0.963018   -2.247  0.026421 *
```

```
## Adult.Mortality          -0.022474    0.003785   -5.938 2.75e-08 ***
## log(percentage.expenditure + 0.5)  0.697748    0.177399    3.933 0.000139 ***
## Total.expenditure          0.227913    0.119144    1.913 0.058083 .
## log(HIV.AIDS)             -2.075600    0.320348   -6.479 2.00e-09 ***
## log(thinness.5.9.years)     -0.614474    0.331445   -1.854 0.066145 .
## Schooling                  0.743394    0.173556    4.283 3.68e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.212 on 123 degrees of freedom
## Multiple R-squared:  0.8682, Adjusted R-squared:  0.8607
## F-statistic: 115.7 on 7 and 123 DF,  p-value: < 2.2e-16
vif(life_exp_2014_lm_mod_new)
```

```
##              Status              Adult.Mortality
##              1.460255              2.189901
## log(percentage.expenditure + 0.5)              Total.expenditure
##              1.619798              1.147883
##              log(HIV.AIDS)              log(thinness.5.9.years)
##              2.307728              1.616289
##              Schooling
##              2.871240
```

Next, we checked which year would fit for our second model. After checking the number of missing values for 2000, 2001, 2002, and 2003, we determined that these years were not the best for our data, since they all left us with less than 100 data points. Therefore, we concluded that 2004 would be much better to use, since it contained less missing values and inaccurate 0 values than the other early years.

```
vec2004 <- c(2004)
LifeExpectancy2004Bad <- LifeExpectancy[LifeExpectancy$Year %in% vec2004, ]
sapply(LifeExpectancy2004Bad, function(x) sum(is.na(x)))
```

```
##              Country              Year
##              0              0
##              Status              Life.expectancy
##              0              0
##              Adult.Mortality              infant.deaths
##              0              0
##              Alcohol              percentage.expenditure
##              1              0
##              Hepatitis.B              Measles
##              45              0
##              BMI              under.five.deaths
##              2              0
##              Polio              Total.expenditure
```

```
##                2                3
##            Diphtheria            HIV.AIDS
##                2                0
##                GDP                Population
##                27                40
##        thinness..1.19.years        thinness.5.9.years
##                2                2
## Income.composition.of.resources        Schooling
##                10                10
```

```
LifeExpectancy2004Old <- na.omit(LifeExpectancy[LifeExpectancy$Year %in% vec2004, ])
sapply(LifeExpectancy2004Old, function(x) sum(x == 0))
```

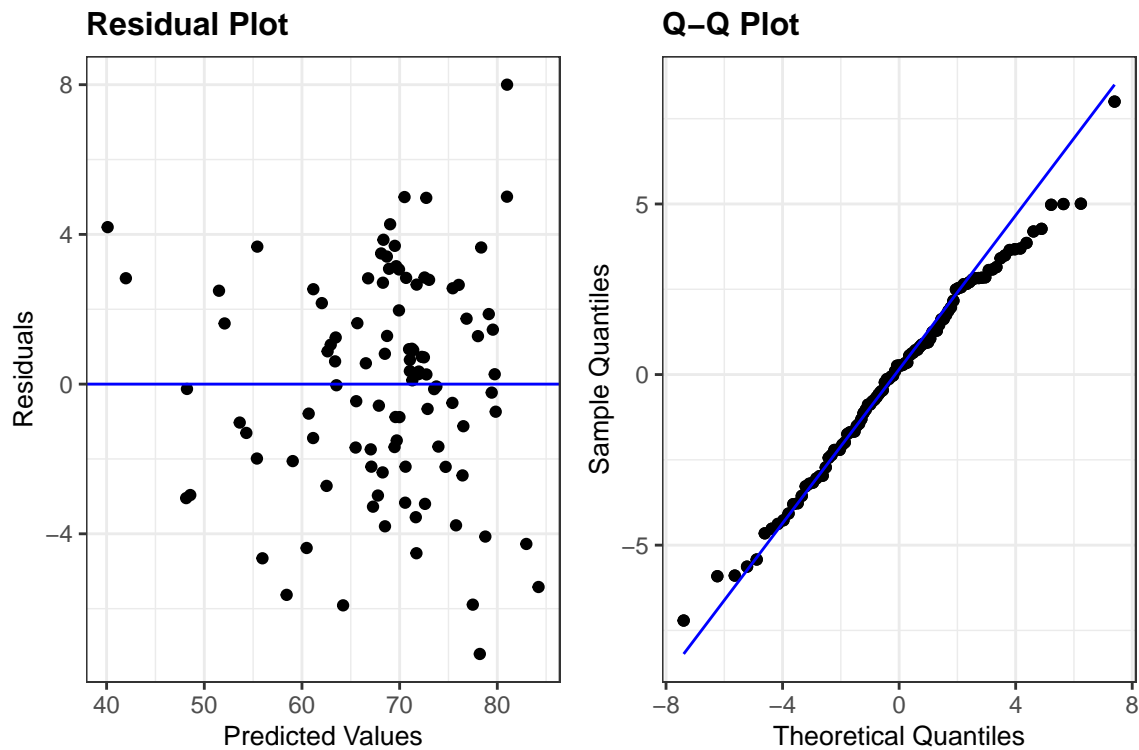
```
##                Country                Year
##                0                0
##                Status                Life.expectancy
##                0                0
##        Adult.Mortality                infant.deaths
##                0                23
##                Alcohol                percentage.expenditure
##                0                1
##                Hepatitis.B                Measles
##                0                33
##                BMI                under.five.deaths
##                0                21
##                Polio                Total.expenditure
##                0                0
##                Diphtheria                HIV.AIDS
##                0                0
##                GDP                Population
##                0                0
##        thinness..1.19.years        thinness.5.9.years
##                0                0
## Income.composition.of.resources        Schooling
##                8                0
```

```
LifeExpectancy2004 <- subset(LifeExpectancy2004Old, select =
                             -c(infant.deaths, Measles, under.five.deaths))
```

Since infant deaths, measles, and under 5 deaths contained a high number of unreasonable 0 values, we decided that it was better to remove these terms altogether, since removing the countries that had these 0 values led to models that were less accurate and contained more variation. Our starting model for 2004 is listed below. Similar to the model for 2014, we checked residual plots and cook's distance to determine if there were any outliers or necessary transformations present in our data. Since many of the terms currently violate the MLR assumption of constant variance, many transformations are needed.

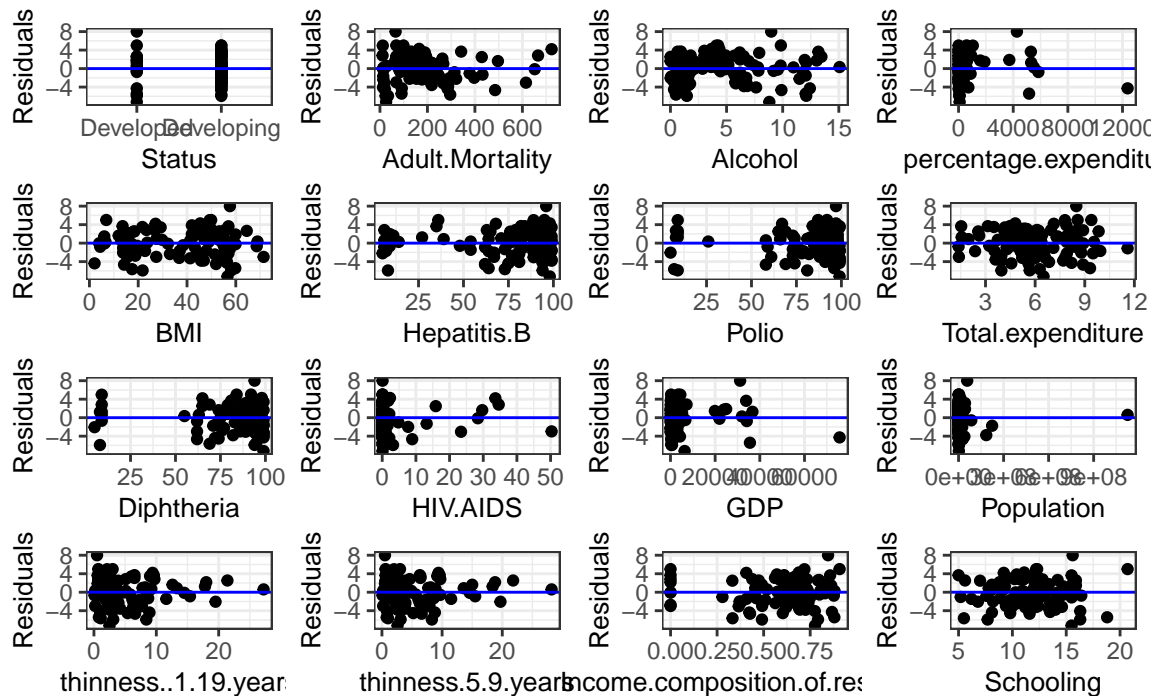
```
LifeExpectancy2004_lm <- lm(Life.expectancy ~ Status + Adult.Mortality +
  Alcohol + percentage.expenditure + BMI +
  Hepatitis.B + Polio + Total.expenditure +
  Diphtheria + HIV.AIDS + GDP + Population +
  thinness..1.19.years + thinness.5.9.years +
  Income.composition.of.resources + Schooling,
  data = LifeExpectancy2004)

resid_panel(LifeExpectancy2004_lm, plots = c("resid", "qq"))
```

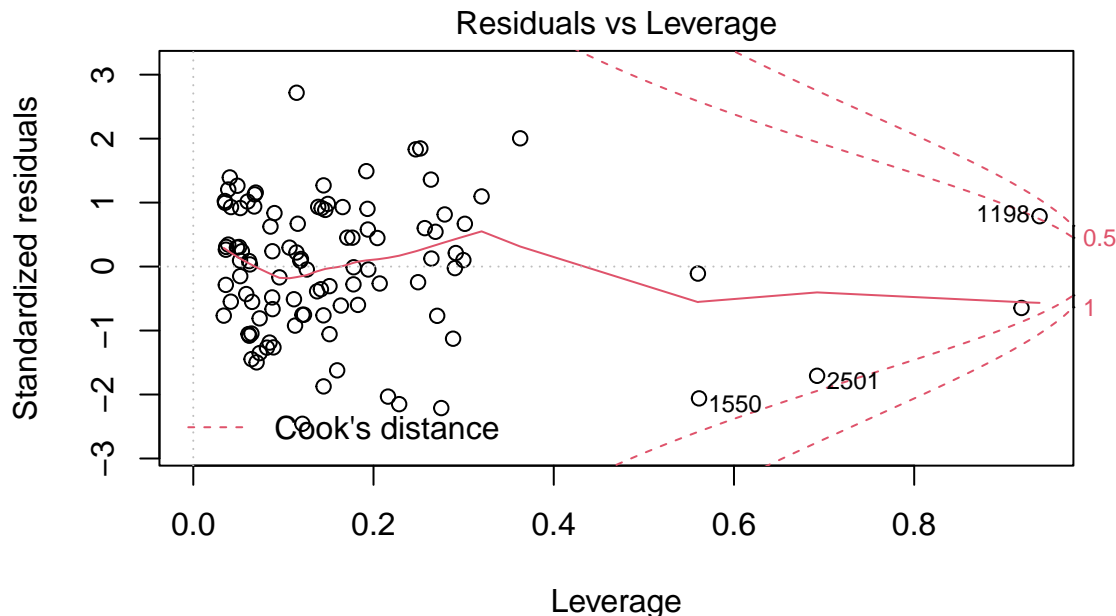


```
resid_xpanel(LifeExpectancy2004_lm)
```

Plots of Residuals vs Predictor Variables



```
plot(LifeExpectancy2004_lm, which = 5)
```

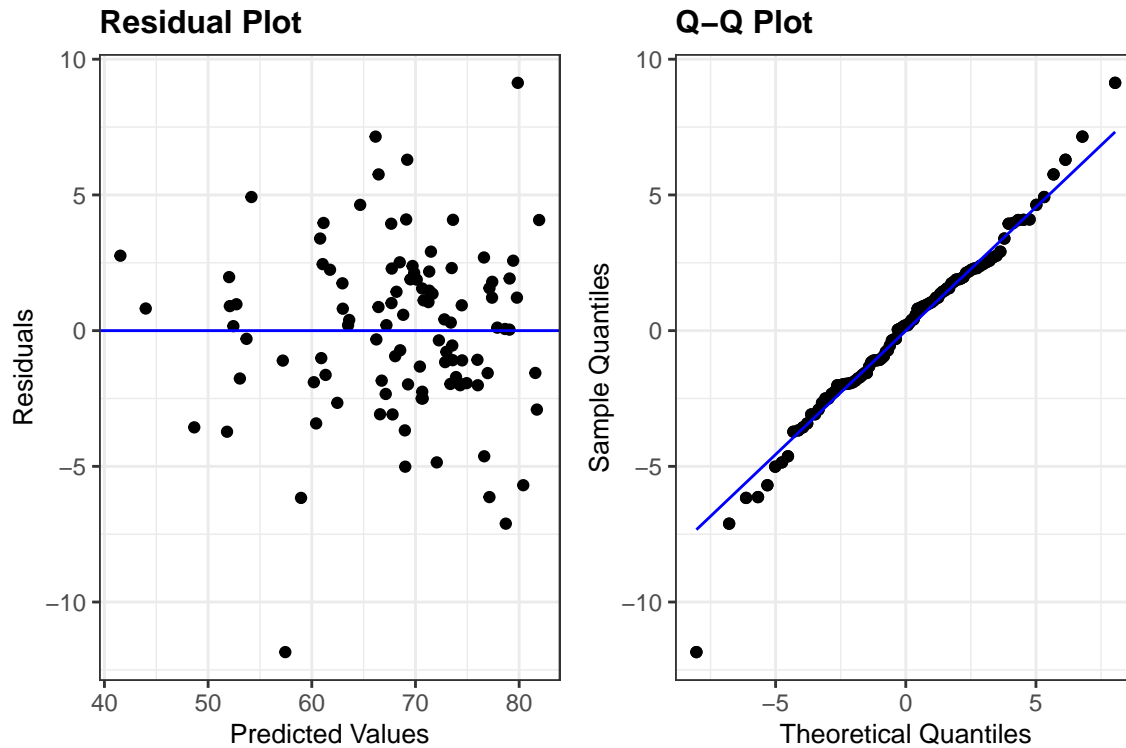


lm(Life.expectancy ~ Status + Adult.Mortality + Alcohol + percentage.expend . After transforming many of the variables, the residual plots look much better, and the MLR assumptions are now met. While some of the countries had higher Cook's distances, none of the points had Cook's distances larger than 1, so we concluded that there were not any influential cases that had to be removed.

```
NewLifeExpectancy2004_lm <- lm(Life.expectancy ~ Status + Adult.Mortality +  
  Alcohol + log(percentage.expenditure+.5) +  
  BMI + I(Hepatitis.B^2) + I(Polio^2) +
```

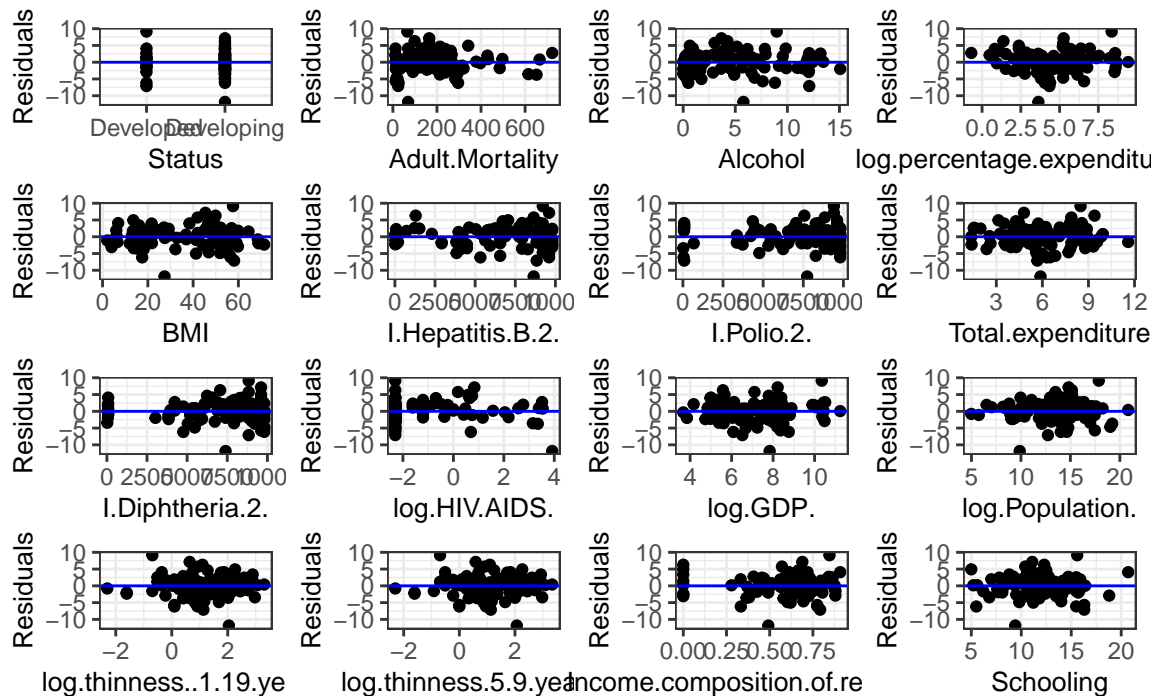
```
Total.expenditure + I(Diphtheria^2) +  
log(HIV.AIDS) + log(GDP) + log(Population) +  
log(thinness..1.19.years) +  
log(thinness.5.9.years) +  
Income.composition.of.resources + Schooling,  
data = LifeExpectancy2004)
```

```
resid_panel(NewLifeExpectancy2004_lm, plots = c("resid", "qq"))
```



```
resid_xpanel(NewLifeExpectancy2004_lm)
```

Plots of Residuals vs Predictor Variables



We now have our transformed model for 2004 listed below.

```
NewLifeExpectancy2004_lm <- lm(Life.expectancy ~ Status + Adult.Mortality +
  Alcohol + log(percentage.expenditure+.5) +
  BMI + I(Hepatitis.B^2) + I(Polio^2) +
  Total.expenditure + I(Diphtheria^2) +
  log(HIV.AIDS) + log(GDP) + log(Population) +
  log(thinness..1.19.years) +
  log(thinness.5.9.years) +
  Income.composition.of.resources + Schooling,
  data = LifeExpectancy2004)

summary(NewLifeExpectancy2004_lm)
```

```
##
## Call:
## lm(formula = Life.expectancy ~ Status + Adult.Mortality + Alcohol +
##   log(percentage.expenditure + 0.5) + BMI + I(Hepatitis.B^2) +
##   I(Polio^2) + Total.expenditure + I(Diphtheria^2) + log(HIV.AIDS) +
##   log(GDP) + log(Population) + log(thinness..1.19.years) +
##   log(thinness.5.9.years) + Income.composition.of.resources +
##   Schooling, data = LifeExpectancy2004)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.844  -1.918   0.201   1.905   9.130
```



```
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      57.4370366   3.8815744  14.797 < 2e-16 ***
## StatusDeveloping -0.2861505   1.3160724  -0.217  0.82839
## Adult.Mortality  -0.0206740   0.0036061  -5.733 1.43e-07 ***
## Alcohol          -0.0603584   0.1462344  -0.413  0.68082
## log(percentage.expenditure + 0.5) 0.5429216   0.4054211   1.339  0.18405
## BMI              -0.0029266   0.0259344  -0.113  0.91041
## I(Hepatitis.B^2) -0.0001434   0.0001396  -1.027  0.30716
## I(Polio^2)        -0.0001828   0.0001645  -1.111  0.26952
## Total.expenditure 0.1114294   0.1732443   0.643  0.52181
## I(Diphtheria^2)   0.0005761   0.0002078   2.773  0.00681 **
## log(HIV.AIDS)     -1.9195525   0.3157504  -6.079 3.22e-08 ***
## log(GDP)          -0.0645592   0.5058171  -0.128  0.89874
## log(Population)    0.0171552   0.1157011   0.148  0.88248
## log(thinness..1.19.years) 1.0496909   1.5664205   0.670  0.50458
## log(thinness.5.9.years) -1.8158104   1.5585206  -1.165  0.24721
## Income.composition.of.resources 4.5476132   2.1982092   2.069  0.04157 *
## Schooling         0.5473130   0.2154090   2.541  0.01285 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.404 on 86 degrees of freedom
## Multiple R-squared:  0.8814, Adjusted R-squared:  0.8594
## F-statistic: 39.96 on 16 and 86 DF,  p-value: < 2.2e-16
```

```
vif(NewLifeExpectancy2004_lm)
```

```
##               Status               Adult.Mortality
##           1.916045                2.407504
##           Alcohol log(percentage.expenditure + 0.5)
##           3.047419                6.255727
##               BMI               I(Hepatitis.B^2)
##           2.035900                1.775367
##           I(Polio^2)           Total.expenditure
##           2.047256                1.201566
##           I(Diphtheria^2)       log(HIV.AIDS)
##           2.832007                2.654211
##           log(GDP)              log(Population)
##           6.385168                1.094936
##           log(thinness..1.19.years) log(thinness.5.9.years)
##           24.490936                25.128522
##           Income.composition.of.resources Schooling
```

```
##                                2.118513                                3.506393
```

After removing the terms that showed high collinearity or were not significant, our final model is shown below, with all terms being statistically significant with 95% confidence. The p-value for Income.Composition.Of.Resources (0.054) we deemed was close enough to warrant not removing it from our model.

```
NewLifeExpectancy2004_lm <- lm(Life.expectancy ~ Adult.Mortality +
                                log(percentage.expenditure+.5) +
                                I(Diphtheria^2) + log(HIV.AIDS) +
                                log(thinness.5.9.years) +
                                Income.composition.of.resources + Schooling,
                                data = LifeExpectancy2004)
summary(NewLifeExpectancy2004_lm)
```

```
##
## Call:
## lm(formula = Life.expectancy ~ Adult.Mortality + log(percentage.expenditure +
##      0.5) + I(Diphtheria^2) + log(HIV.AIDS) + log(thinness.5.9.years) +
##      Income.composition.of.resources + Schooling, data = LifeExpectancy2004)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.6853  -1.8300   0.1567   1.9256   9.2102
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      57.9487767   2.2412183   25.856 < 2e-16 ***
## Adult.Mortality    -0.0210815   0.0034408   -6.127 2.02e-08 ***
## log(percentage.expenditure + 0.5)  0.4373682   0.2091528    2.091 0.03918 *
## I(Diphtheria^2)     0.0003384   0.0001338    2.529 0.01308 *
## log(HIV.AIDS)      -1.9052873   0.2877490   -6.621 2.11e-09 ***
## log(thinness.5.9.years) -0.8520259   0.3911610   -2.178 0.03187 *
## Income.composition.of.resources    3.9737010   2.0384236    1.949 0.05420 .
## Schooling           0.5286524   0.1854588    2.851 0.00535 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.305 on 95 degrees of freedom
## Multiple R-squared:  0.8765, Adjusted R-squared:  0.8674
## F-statistic: 96.34 on 7 and 95 DF,  p-value: < 2.2e-16
vif(NewLifeExpectancy2004_lm)
```

```
##              Adult.Mortality log(percentage.expenditure + 0.5)
##              2.324657                                1.765774
```

##	I(Diphtheria^2)	log(HIV.AIDS)
##	1.245397	2.337858
##	log(thinness.5.9.years)	Income.composition.of.resources
##	1.678788	1.932078
##	Schooling	
##	2.756582	