

STAT 346 Final Project

Can Money Buy Happiness?: Understanding the Differential Between a Country's GDP Rank and Happiness Rank

Edith Edwards-Mizel, Varya Kluev, Greta Laesch

Adding and Cleaning Data

```
# Explanatory Variables
census_data <- read_csv("~/Desktop/export.csv", show_col_types = FALSE)
who_data <- read_csv("~/Desktop/Annex 2-4-Table 1.csv", show_col_types = FALSE)
gdp_data <- read_csv("~/Desktop/GDP_data.csv", show_col_types = FALSE)
wh_data <- read_csv("~/Desktop/WHO.csv", show_col_types = FALSE)
who_data <- who_data[-(1:4),]
colnames(census_data)[4] = "country"
colnames(who_data)[1] = "country"
colnames(who_data)[2] = "Clean_energy_access_prop"
colnames(who_data)[3] = "Harmful_air_mean_conc"
colnames(who_data)[4] = "Mortality_homicide_rate"
colnames(who_data)[5] = "Adult_obesity_rate"
colnames(who_data)[6] = "Tobacco_use_rate"
colnames(who_data)[7] = "Expenditure_health_perc"
colnames(who_data)[8] = "Mortality_suicide_rate"
colnames(who_data)[9] = "Alcohol_consumption_liters"
df_merge <- merge(census_data, who_data, by = "country")
df_merge <- df_merge[-c(2:5) ]

# Categorical Variable
hc_data <- read_csv("~/Desktop/UNHC.csv", show_col_types = FALSE)
univ <- hc_data$country
univ_var <- c()
for (i in 1:nrow(df_merge)) {
  ifelse(df_merge[i,1] %in% univ, univ_var[i] <- 1, univ_var[i] <- 0)
}
df_merge["Univ_hc"] <- univ_var
df_merge$Univ_hc <- as.factor(df_merge$Univ_hc)

# Covid Deaths
covid_deaths <- read_csv("~/Desktop/covidData.csv", show_col_types = FALSE)
# -- Cumulative Deaths from 01/03/2020-12/31/2022
newest_covid <- covid_deaths[covid_deaths$Date_reported == "2022-12-31", ]
cv_dummy_df <- data.frame(newest_covid$Country, newest_covid$Cumulative_deaths)
colnames(cv_dummy_df)[1] <- "country"
colnames(cv_dummy_df)[2] <- "tot_deaths"
df_merge <- merge(df_merge, cv_dummy_df, by = "country")

# Corruption
```

```

cpi_data <- read_csv("~/Desktop/CPI.csv", show_col_types = FALSE)
# -- Lower Scores == More Corruption
cpi_data <- cpi_data[-(1:2),]
cpi_dummy_df <- data.frame(cpi_data$`Corruption Perceptions Index 2022: Global scores`, cpi_data$...4)
colnames(cpi_dummy_df)[1] <- "country"
colnames(cpi_dummy_df)[2] <- "cpi_score"
df_merge <- merge(df_merge, cpi_dummy_df, by = "country")
df_merge$cpi_score <- as.numeric(df_merge$cpi_score)

# Response Variable
gdp_dummy_df <- data.frame(gdp_data$country, gdp_data$rank)
colnames(gdp_dummy_df)[1] <- "country"
colnames(gdp_dummy_df)[2] <- "GDP_rank"

df_merge <- merge(df_merge, gdp_dummy_df, by = "country")

wh_dummy_df <- data.frame(wh_data$country, wh_data$rank)
colnames(wh_dummy_df)[1] <- "country"
colnames(wh_dummy_df)[2] <- "WH_rank"

df_merge <- merge(df_merge, wh_dummy_df, by = "country")

df_merge$GDP_rank_std <- ((df_merge$GDP_rank) - mean(df_merge$GDP_rank))/sd(df_merge$GDP_rank)
df_merge$WH_rank_std <- ((df_merge$WH_rank) - mean(df_merge$WH_rank))/sd(df_merge$WH_rank)

df_merge["DIF_rank"] <- (df_merge$GDP_rank_std - df_merge$WH_rank_std)

# Cleaning Data
# Make missing data NA
for (i in 1:length(df_merge)) { # columns
  for (j in 1:nrow(df_merge)) { # rows
    ifelse(df_merge[j,i] == "-", df_merge[j,i] <- NA, NA)
  }
}

# Round <1 to 0
for (i in 1:length(df_merge)) { # columns
  for (j in 1:nrow(df_merge)) { # rows
    ifelse(df_merge[j,i] == "<1" || df_merge[j,i] == "<0.1", df_merge[j,i] <- 0, NA)
  }
}

# Convert Characters to Numeric
for (i in (8:15)) {
  df_merge[,i] <- as.numeric(df_merge[,i])
}

# Rename columns with spaces
colnames(df_merge)[2:7] <- c("Population", "rate_nat_increase", "tot_fert_rate", "life_exp_males", "lif
attach(df_merge)

```

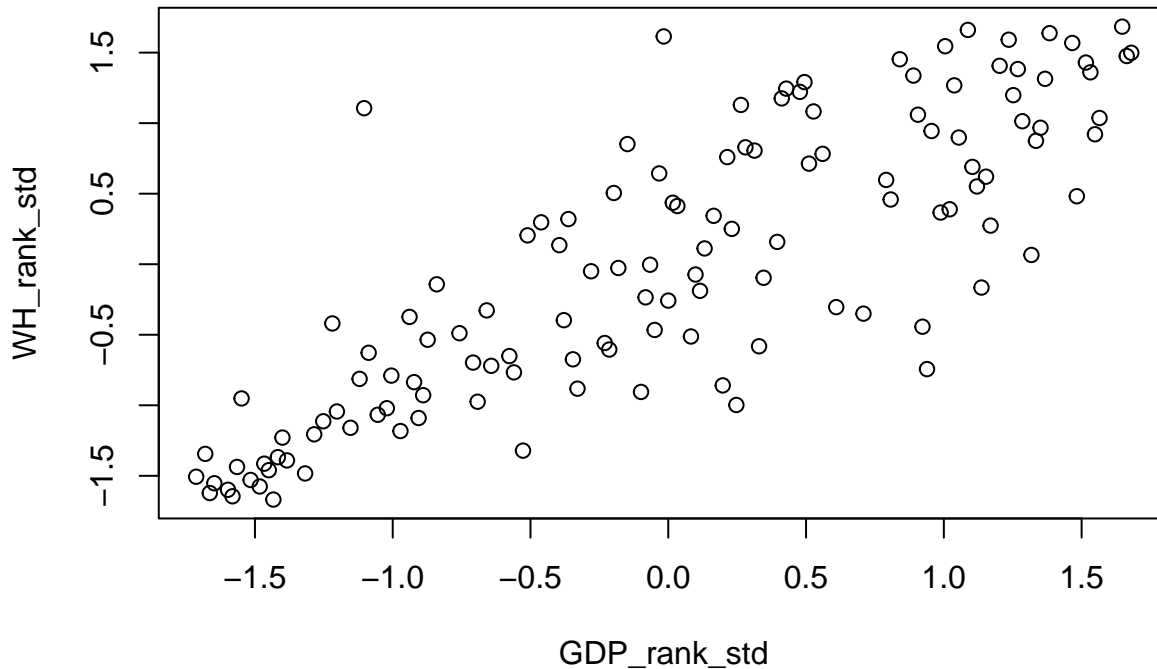
```
#write.csv(df_merge, "/Users/Greta/Desktop/master_df.csv", row.names=TRUE)
```

Goal: Understanding the Phenomenon in General

Exploratory Data Analysis

```
# Response Variables
```

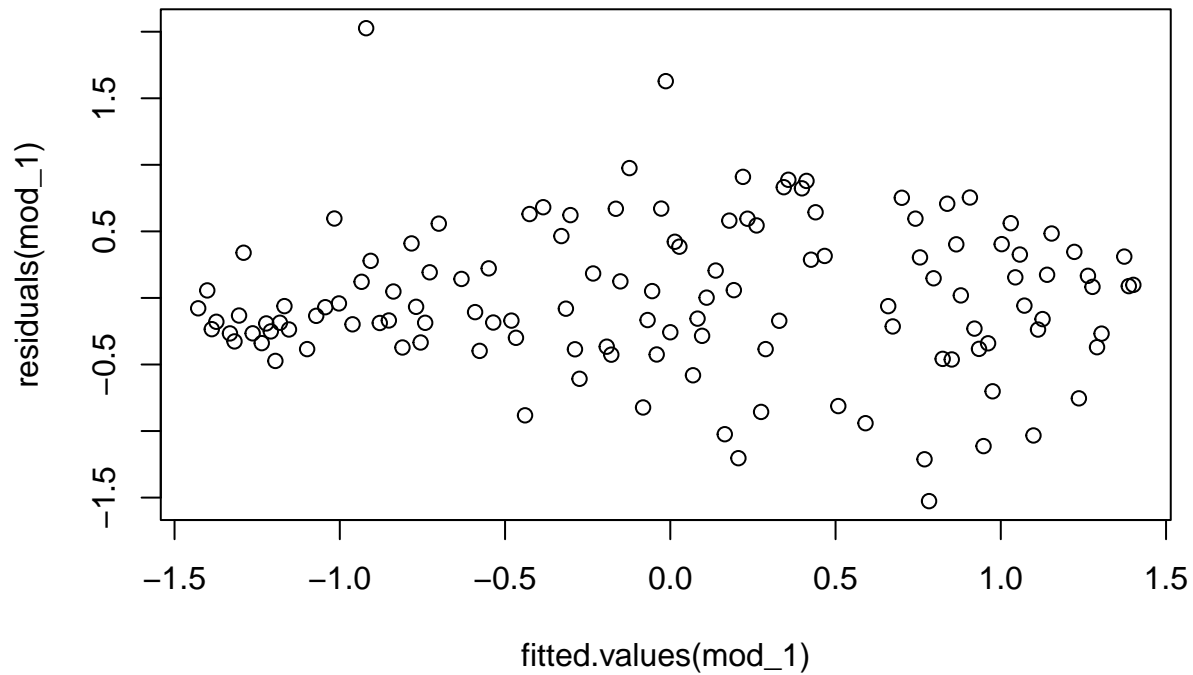
```
plot(WH_rank_std ~ GDP_rank_std)
```



```
mod_1 <- lm(WH_rank_std~GDP_rank_std)
summary(mod_1)
```

```
##
## Call:
## lm(formula = WH_rank_std ~ GDP_rank_std)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.52639 -0.31305 -0.06676  0.34281  2.02637
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.576e-16  4.918e-02   0.00    1
## GDP_rank_std  8.338e-01  4.938e-02  16.89 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5543 on 125 degrees of freedom
## Multiple R-squared:  0.6952, Adjusted R-squared:  0.6928
## F-statistic: 285.2 on 1 and 125 DF, p-value: < 2.2e-16
```

```
res1 <- plot(residuals(mod_1) ~ fitted.values(mod_1))
```



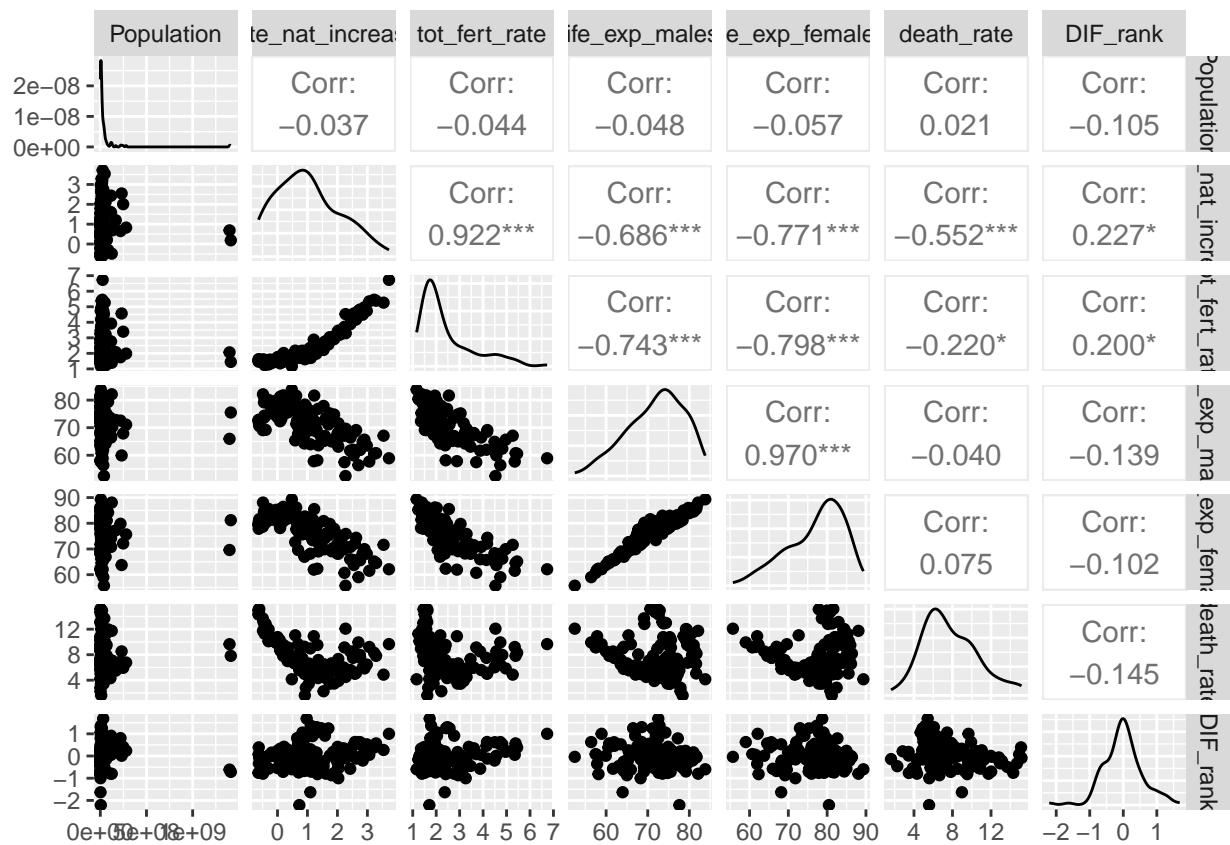
```
res1
```

```
## NULL
```

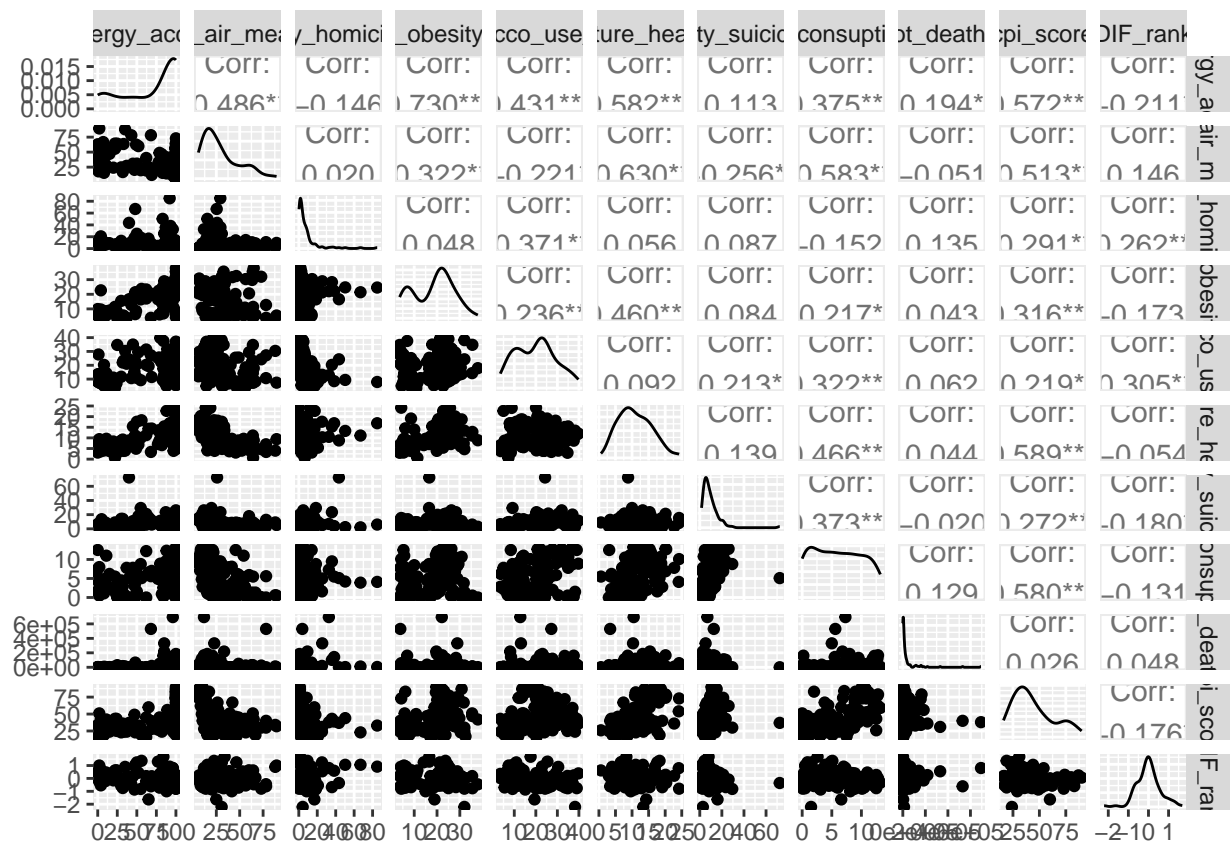
```
# Explanation: We can see that GDP explains some, but not all of variation in the happiness rating  
# of a country, so what else is there to explain why some countries are happier than others,  
# compared to
```

```
# Census Data Pairs
```

```
ggpairs(df_merge[,c(2,3,4,5,6,7,23)])
```



```
# WHO Data Pairs
ggpairs(df_merge[,c(8,9,10,11,12,13,14,15,17,18,23)])
```

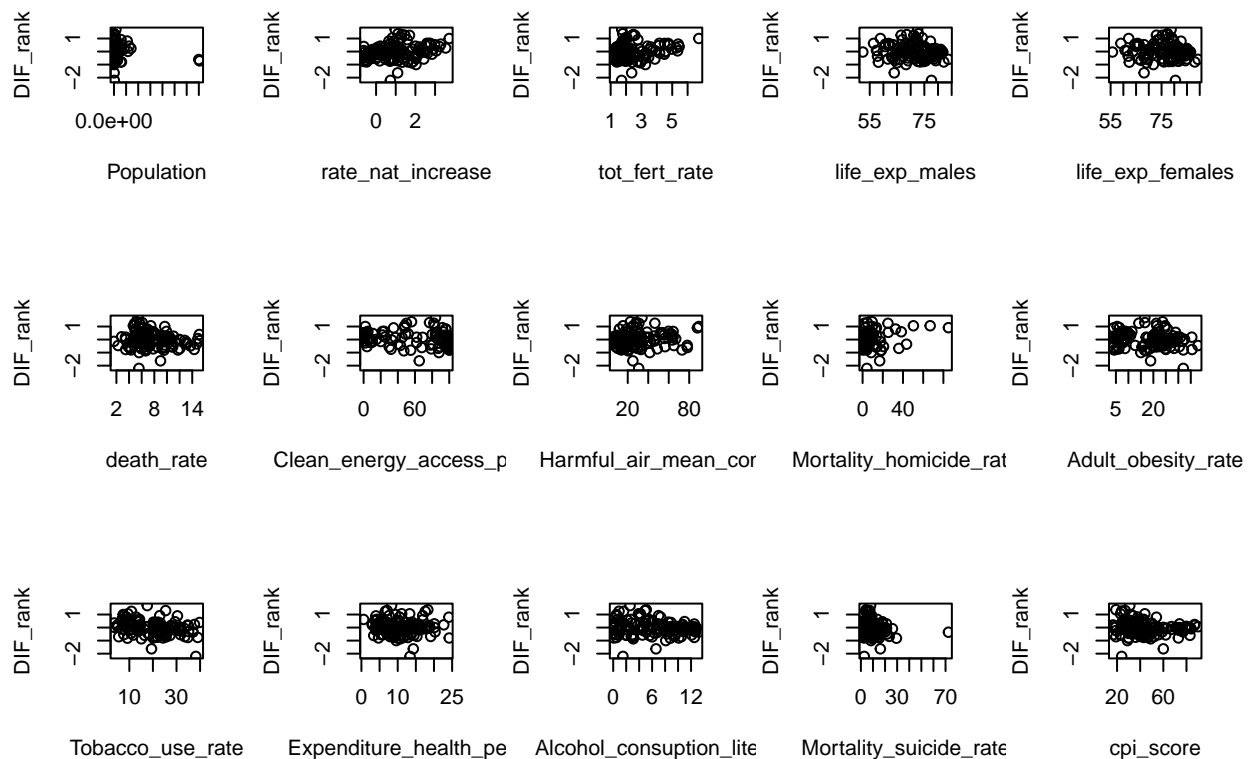


Transformations

```

par(mfrow=c(3,5))
plot(DIF_rank~Population)
plot(DIF_rank~rate_nat_increase)
plot(DIF_rank~tot_fert_rate)
plot(DIF_rank~life_exp_males)
plot(DIF_rank~life_exp_females)
plot(DIF_rank~death_rate)
plot(DIF_rank~Clean_energy_access_prop)
plot(DIF_rank~Harmful_air_mean_conc)
plot(DIF_rank~Mortality_homicide_rate)
plot(DIF_rank~Adult_obesity_rate)
plot(DIF_rank~Tobacco_use_rate)
plot(DIF_rank~Expenditure_health_perc)
plot(DIF_rank~Alcohol_consumption_liters)
plot(DIF_rank~Mortality_suicide_rate)
plot(DIF_rank~cpi_score)

```



```
# Transform Non-Linear Relationships as Necessary
df_merge$PopulationT <- log(df_merge$Population)
df_merge$tot_fert_rateT <- log(df_merge$tot_fert_rate)
df_merge$Harmful_air_mean_concT <- log(df_merge$Harmful_air_mean_conc)
df_merge$Mortality_homicide_rateT <- log(df_merge$Mortality_homicide_rate)
df_merge$Mortality_suicide_rateT <- log(df_merge$Mortality_suicide_rate)
df_merge$cpi_scoreT <- log(df_merge$cpi_score)
```

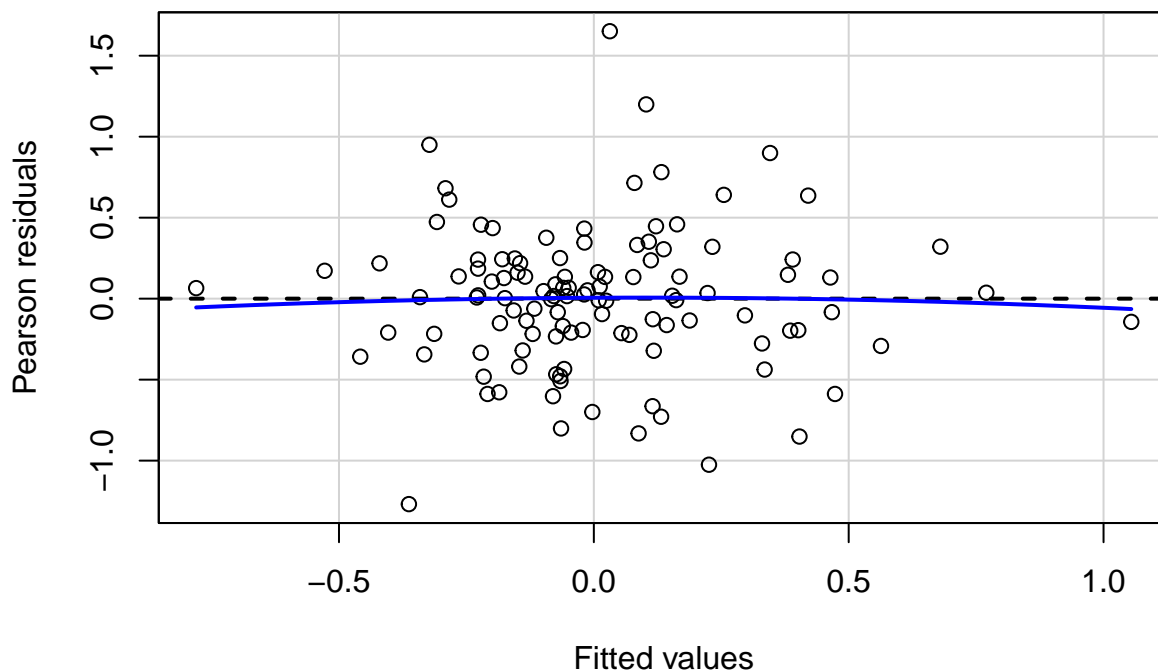
Preliminary Data Analysis: Model Selection Using Automatic Selection Procedure

```
df_merge_noNA <- na.omit(df_merge)
# Original Model Before Transformations
myformula <- formula(paste(colnames(df_merge)[23], paste(colnames(df_merge_noNA)[2:18], collapse = "+"))
mod_noT <- lm(myformula, data = df_merge)
summary(mod_noT)

##
## Call:
## lm(formula = myformula, data = df_merge)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26838 -0.21745  0.01623  0.22782  1.65104
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.910e+00  2.412e+00  -1.206   0.2306
## Population    -6.908e-10  3.147e-10  -2.195   0.0306 *
## rate_nat_increase  3.397e-02  6.412e-01   0.053   0.9579
## tot_fert_rate    7.263e-02  3.962e-01   0.183   0.8549
```

```
## life_exp_males          -2.915e-02  3.716e-02 -0.784  0.4347
## life_exp_females        6.340e-02  3.835e-02  1.653  0.1015
## death_rate              1.136e-02  9.748e-02  0.117  0.9075
## Clean_energy_access_prop -1.476e-03  3.270e-03 -0.452  0.6526
## Harmful_air_mean_conc    6.946e-03  3.506e-03  1.981  0.0504 .
## Mortality_homicide_rate  1.105e-02  5.160e-03  2.142  0.0347 *
## Adult_obesity_rate      -1.230e-02  8.978e-03 -1.370  0.1740
## Tobacco_use_rate        -3.274e-03  7.482e-03 -0.438  0.6626
## Expenditure_health_perc  1.106e-02  1.509e-02  0.733  0.4654
## Mortality_suicide_rate  -8.532e-03  7.287e-03 -1.171  0.2446
## Alcohol_consuption_liters 7.190e-04  1.883e-02  0.038  0.9696
## Univ_hc1                -1.609e-01  1.250e-01 -1.287  0.2011
## tot_deaths              1.160e-06  6.207e-07  1.869  0.0647 .
## cpi_score               -9.258e-04  4.724e-03 -0.196  0.8450
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4749 on 97 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.276, Adjusted R-squared:  0.1491
## F-statistic: 2.175 on 17 and 97 DF, p-value: 0.009317
```

```
residualPlot(mod_noT)
```



```
# Model with Transformed Non-Linear Variables
```

```
mod_T <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_females +
summary(mod_T)
```

```
##
```

```
## Call:
```

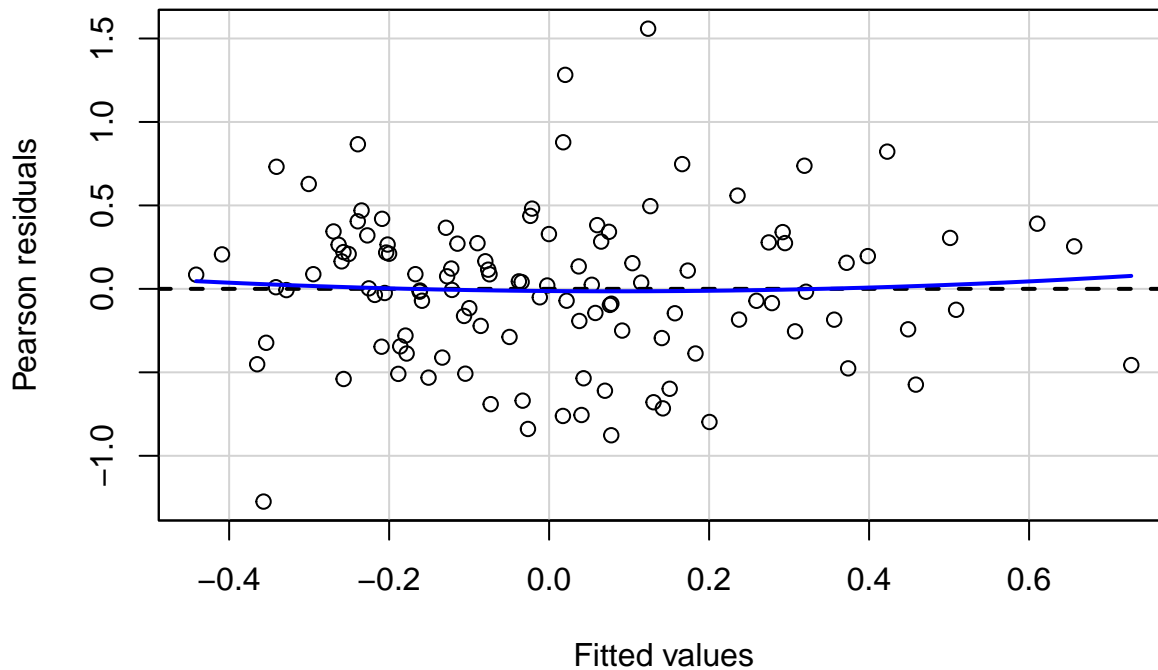
```
## lm(formula = DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##     life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##     Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
##     Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
```



```

##      Alcohol_consumption_liters + Univ_hc + tot_deaths + cpi_score,
##      data = df_merge)
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.27421 -0.28397  0.00993  0.27232  1.55870
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -5.262e+00  2.010e+00  -2.618  0.0103 *
## PopulationT        3.006e-02  4.184e-02   0.718  0.4742
## rate_nat_increase   6.349e-01  3.747e-01   1.695  0.0934 .
## tot_fert_rateT     -9.557e-01  7.416e-01  -1.289  0.2006
## life_exp_males    -2.156e-02  4.191e-02  -0.514  0.6081
## life_exp_females    7.412e-02  3.664e-02   2.023  0.0458 *
## death_rate        8.809e-02  5.113e-02   1.723  0.0881 .
## Clean_energy_access_prop -2.275e-03  3.447e-03  -0.660  0.5108
## Harmful_air_mean_concT  9.643e-02  1.147e-01   0.841  0.4024
## Mortality_homicide_rateT 1.155e-01  7.218e-02   1.600  0.1128
## Adult_obesity_rate  -3.156e-03  9.702e-03  -0.325  0.7457
## Tobacco_use_rate   -4.480e-03  7.696e-03  -0.582  0.5619
## Expenditure_health_perc  8.109e-03  1.603e-02   0.506  0.6140
## Mortality_suicide_rateT -5.618e-02  1.005e-01  -0.559  0.5775
## Alcohol_consumption_liters -5.261e-03  1.935e-02  -0.272  0.7862
## Univ_hc1         -1.736e-01  1.341e-01  -1.295  0.1983
## tot_deaths        3.856e-07  6.802e-07   0.567  0.5721
## cpi_score         5.148e-04  5.218e-03   0.099  0.9216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.491 on 97 degrees of freedom
## (12 observations deleted due to missingness)
## Multiple R-squared:  0.2259, Adjusted R-squared:  0.0902
## F-statistic: 1.665 on 17 and 97 DF, p-value: 0.06288
residualPlot(mod_T)

```



```
# Model Selection
mod_noNA <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop + Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate + Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT + Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT)
# Summary matches model with missing data
bw_aic2 = step(mod_noNA, direction = "backward")
```

```
## Start: AIC=-147.18
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
##   Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##   Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT
##
##
```

	Df	Sum of Sq	RSS	AIC
## - cpi_scoreT	1	0.00469	23.389	-149.16
## - Alcohol_consuption_liters	1	0.01852	23.403	-149.09
## - Adult_obesity_rate	1	0.02523	23.410	-149.06
## - Expenditure_health_perc	1	0.05960	23.444	-148.89
## - life_exp_males	1	0.06658	23.451	-148.85
## - tot_deaths	1	0.07602	23.460	-148.81
## - Mortality_suicide_rateT	1	0.08018	23.465	-148.79
## - Tobacco_use_rate	1	0.08366	23.468	-148.77
## - Clean_energy_access_prop	1	0.10752	23.492	-148.65
## - PopulationT	1	0.12706	23.511	-148.56
## - Harmful_air_mean_concT	1	0.16777	23.552	-148.36
## - tot_fert_rateT	1	0.40275	23.787	-147.22
## - Univ_hc	1	0.40953	23.794	-147.18
## <none>			23.384	-147.18
## - Mortality_homicide_rateT	1	0.63071	24.015	-146.12
## - rate_nat_increase	1	0.69035	24.075	-145.83
## - death_rate	1	0.71009	24.095	-145.74
## - life_exp_females	1	0.99496	24.379	-144.39
##				

```
## Step: AIC=-149.16
```

```

## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
##   Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##   Alcohol_consuption_liters + Univ_hc + tot_deaths
##
##
##      Df Sum of Sq  RSS   AIC
## - Alcohol_consuption_liters  1   0.01663 23.406 -151.07
## - Adult_obesity_rate         1   0.02717 23.416 -151.02
## - life_exp_males             1   0.06193 23.451 -150.85
## - Expenditure_health_perc    1   0.06740 23.456 -150.82
## - Mortality_suicide_rateT    1   0.07605 23.465 -150.78
## - tot_deaths                 1   0.07809 23.467 -150.77
## - Tobacco_use_rate           1   0.08421 23.473 -150.74
## - Clean_energy_access_prop    1   0.10288 23.492 -150.65
## - PopulationT                1   0.12239 23.512 -150.56
## - Harmful_air_mean_concT     1   0.16822 23.557 -150.33
## - tot_fert_rateT            1   0.39849 23.788 -149.21
## <none>                       23.389 -149.16
## - Univ_hc                    1   0.42095 23.810 -149.10
## - Mortality_homicide_rateT   1   0.62748 24.017 -148.11
## - rate_nat_increase          1   0.69429 24.083 -147.79
## - death_rate                 1   0.73347 24.122 -147.60
## - life_exp_females           1   0.99154 24.381 -146.38
##
## Step:  AIC=-151.07
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
##   Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##   Univ_hc + tot_deaths
##
##
##      Df Sum of Sq  RSS   AIC
## - Adult_obesity_rate         1   0.03293 23.439 -152.91
## - life_exp_males             1   0.05668 23.462 -152.80
## - Expenditure_health_perc    1   0.05919 23.465 -152.78
## - tot_deaths                 1   0.07020 23.476 -152.73
## - Tobacco_use_rate           1   0.08617 23.492 -152.65
## - Clean_energy_access_prop    1   0.09034 23.496 -152.63
## - Mortality_suicide_rateT    1   0.09421 23.500 -152.61
## - PopulationT                1   0.13040 23.536 -152.44
## - Harmful_air_mean_concT     1   0.19365 23.599 -152.13
## - tot_fert_rateT            1   0.40286 23.809 -151.11
## <none>                       23.406 -151.07
## - Univ_hc                    1   0.44108 23.847 -150.93
## - Mortality_homicide_rateT   1   0.64134 24.047 -149.97
## - rate_nat_increase          1   0.68914 24.095 -149.74
## - death_rate                 1   0.72763 24.133 -149.55
## - life_exp_females           1   0.98582 24.392 -148.33
##
## Step:  AIC=-152.91
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +

```

```

##      Expenditure_health_perc + Mortality_suicide_rateT + Univ_hc +
##      tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - Expenditure_health_perc    1    0.06344  23.502 -154.60
## - tot_deaths                  1    0.06862  23.507 -154.58
## - Tobacco_use_rate            1    0.07925  23.518 -154.53
## - life_exp_males              1    0.08701  23.526 -154.49
## - Mortality_suicide_rateT     1    0.10968  23.548 -154.38
## - PopulationT                 1    0.17077  23.609 -154.08
## - Harmful_air_mean_concT      1    0.17664  23.615 -154.05
## - Clean_energy_access_prop    1    0.26990  23.709 -153.60
## <none>                        23.439 -152.91
## - Univ_hc                     1    0.41298  23.852 -152.90
## - tot_fert_rateT              1    0.51665  23.955 -152.41
## - Mortality_homicide_rateT    1    0.61300  24.052 -151.94
## - death_rate                  1    0.77897  24.218 -151.15
## - rate_nat_increase           1    0.79606  24.235 -151.07
## - life_exp_females            1    1.10320  24.542 -149.62
##
## Step: AIC=-154.6
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Mortality_suicide_rateT + Univ_hc + tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - tot_deaths                  1    0.05047  23.552 -156.35
## - life_exp_males              1    0.05627  23.558 -156.33
## - Mortality_suicide_rateT     1    0.09559  23.598 -156.13
## - Tobacco_use_rate            1    0.11542  23.617 -156.04
## - Harmful_air_mean_concT      1    0.13296  23.635 -155.95
## - PopulationT                 1    0.16562  23.668 -155.79
## - Clean_energy_access_prop    1    0.23243  23.735 -155.47
## - Univ_hc                     1    0.36543  23.867 -154.83
## <none>                        23.502 -154.60
## - tot_fert_rateT              1    0.56364  24.066 -153.88
## - rate_nat_increase           1    0.85506  24.357 -152.49
## - Mortality_homicide_rateT    1    0.89266  24.395 -152.31
## - death_rate                  1    0.91665  24.419 -152.20
## - life_exp_females            1    1.06634  24.568 -151.50
##
## Step: AIC=-156.36
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Mortality_suicide_rateT + Univ_hc
##
##              Df Sum of Sq    RSS    AIC
## - life_exp_males              1    0.06535  23.618 -158.04
## - Mortality_suicide_rateT     1    0.10036  23.653 -157.87
## - Tobacco_use_rate            1    0.11740  23.670 -157.78
## - Harmful_air_mean_concT      1    0.12177  23.674 -157.76
## - Clean_energy_access_prop    1    0.21723  23.770 -157.30

```

```

## - Univ_hc                1    0.32863 23.881 -156.76
## - PopulationT            1    0.36906 23.922 -156.57
## <none>                    23.552 -156.35
## - tot_fert_rateT         1    0.62552 24.178 -155.34
## - rate_nat_increase      1    0.88726 24.440 -154.10
## - death_rate             1    0.96002 24.513 -153.76
## - Mortality_homicide_rateT 1    1.06320 24.616 -153.28
## - life_exp_females       1    1.09900 24.651 -153.11
##
## Step:  AIC=-158.04
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Mortality_suicide_rateT + Univ_hc
##
##              Df Sum of Sq    RSS    AIC
## - Mortality_suicide_rateT  1    0.08963 23.707 -159.60
## - Tobacco_use_rate        1    0.11271 23.731 -159.49
## - Harmful_air_mean_concT   1    0.15192 23.770 -159.30
## - Clean_energy_access_prop  1    0.24578 23.864 -158.85
## - PopulationT             1    0.37924 23.997 -158.21
## <none>                     23.618 -158.04
## - Univ_hc                 1    0.45166 24.070 -157.86
## - tot_fert_rateT          1    0.63552 24.253 -156.98
## - rate_nat_increase       1    0.88242 24.500 -155.82
## - death_rate              1    1.15054 24.768 -154.57
## - Mortality_homicide_rateT 1    1.55533 25.173 -152.70
## - life_exp_females        1    2.72589 26.344 -147.47
##
## Step:  AIC=-159.6
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Univ_hc
##
##              Df Sum of Sq    RSS    AIC
## - Tobacco_use_rate        1    0.09059 23.798 -161.16
## - Harmful_air_mean_concT   1    0.19643 23.904 -160.65
## - Clean_energy_access_prop  1    0.20466 23.912 -160.61
## <none>                     23.707 -159.60
## - PopulationT             1    0.42381 24.131 -159.56
## - Univ_hc                 1    0.43364 24.141 -159.52
## - tot_fert_rateT          1    0.65127 24.359 -158.49
## - rate_nat_increase       1    0.96322 24.671 -157.02
## - death_rate              1    1.09977 24.807 -156.39
## - Mortality_homicide_rateT 1    1.66720 25.375 -153.79
## - life_exp_females        1    2.90017 26.608 -148.33
##
## Step:  AIC=-161.16
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Univ_hc
##
##              Df Sum of Sq    RSS    AIC

```

```
## - Harmful_air_mean_concT      1    0.14871 23.947 -162.45
## - Clean_energy_access_prop    1    0.18448 23.983 -162.27
## - Univ_hc                     1    0.41310 24.211 -161.18
## <none>                        23.798 -161.16
## - PopulationT                 1    0.43331 24.231 -161.09
## - tot_fert_rateT              1    0.76181 24.560 -159.54
## - death_rate                  1    1.19890 24.997 -157.51
## - rate_nat_increase           1    1.27895 25.077 -157.14
## - Mortality_homicide_rateT    1    2.04968 25.848 -153.66
## - life_exp_females            1    3.12127 26.919 -148.99
##
## Step:  AIC=-162.45
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Univ_hc
```

	Df	Sum of Sq	RSS	AIC
## - Clean_energy_access_prop	1	0.16184	24.109	-163.67
## <none>			23.947	-162.45
## - Univ_hc	1	0.59035	24.537	-161.65
## - PopulationT	1	0.72955	24.676	-161.00
## - tot_fert_rateT	1	0.75942	24.706	-160.86
## - death_rate	1	1.08901	25.036	-159.33
## - rate_nat_increase	1	1.26635	25.213	-158.52
## - Mortality_homicide_rateT	1	1.91092	25.858	-155.62
## - life_exp_females	1	3.08982	27.037	-150.49

```
## Step:  AIC=-163.67
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Mortality_homicide_rateT +
##      Univ_hc
```

	Df	Sum of Sq	RSS	AIC
## <none>			24.109	-163.67
## - Univ_hc	1	0.72546	24.834	-162.26
## - PopulationT	1	0.78115	24.890	-162.00
## - tot_fert_rateT	1	0.82082	24.930	-161.82
## - death_rate	1	1.33312	25.442	-159.48
## - rate_nat_increase	1	1.60621	25.715	-158.25
## - Mortality_homicide_rateT	1	1.78209	25.891	-157.47
## - life_exp_females	1	2.92880	27.038	-152.49

```
# AIC-Determined Model
```

```
# ----- So different depending on using transformed or not
```

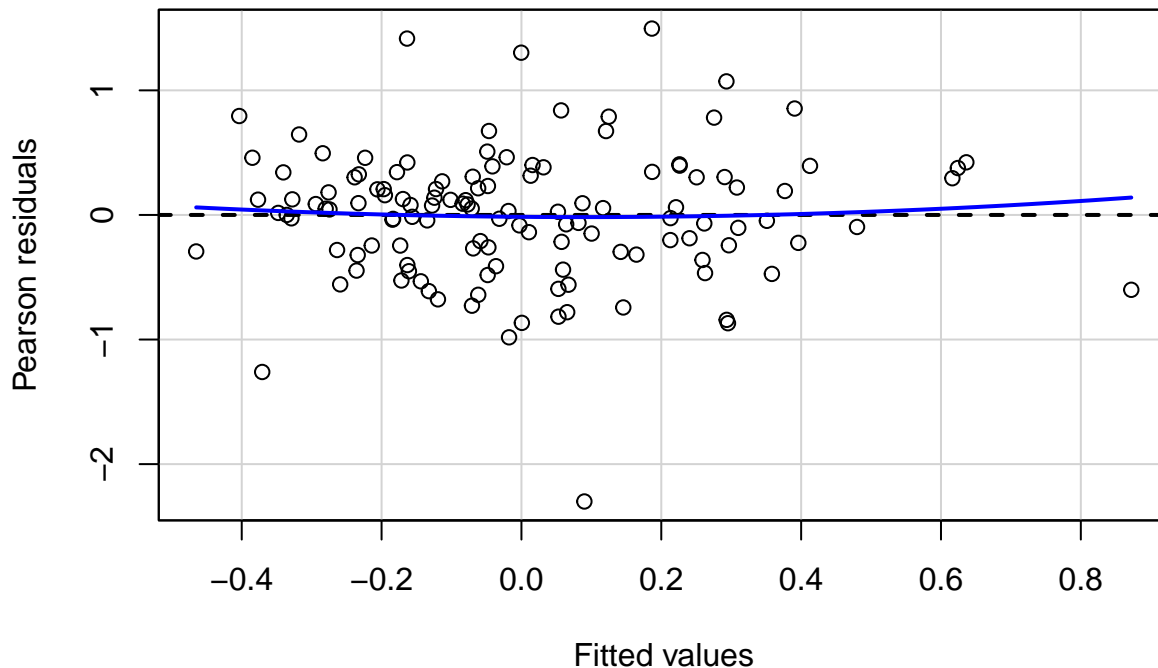
```
mod_bwAIC <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_females + death_rate)
summary(mod_bwAIC)
```

```
##
## Call:
## lm(formula = DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_females + death_rate + Mortality_homicide_rateT +
##      Univ_hc, data = df_merge)
##
## Residuals:
```

##	Min	1Q	Median	3Q	Max
----	-----	----	--------	----	-----

```
## -2.29981 -0.28724 0.04438 0.30543 1.49585
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -5.76174     1.62846   -3.538 0.000576 ***
## PopulationT      0.06939     0.03410    2.035 0.044056 *
## rate_nat_increase 0.89137     0.33432    2.666 0.008738 **
## tot_fert_rateT   -1.34188     0.68193   -1.968 0.051422 .
## life_exp_females  0.05077     0.01579    3.215 0.001677 **
## death_rate       0.10354     0.04371    2.369 0.019463 *
## Mortality_homicide_rateT 0.12909     0.05106    2.528 0.012775 *
## Univ_hc1        -0.20644     0.11550   -1.787 0.076442 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5367 on 119 degrees of freedom
## Multiple R-squared:  0.1814, Adjusted R-squared:  0.1333
## F-statistic: 3.768 on 7 and 119 DF,  p-value: 0.001003
```

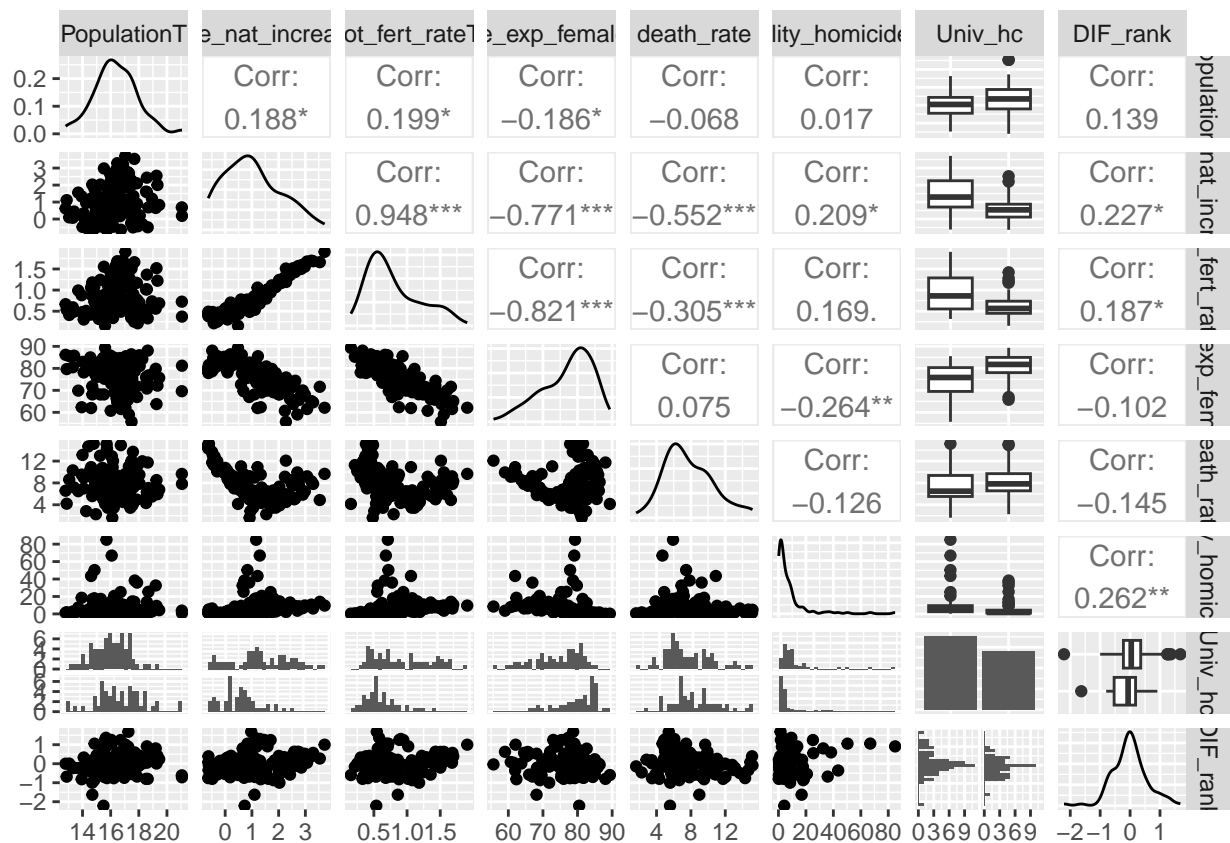
```
residualPlot(mod_bwAIC)
```



```
# Final Model
```

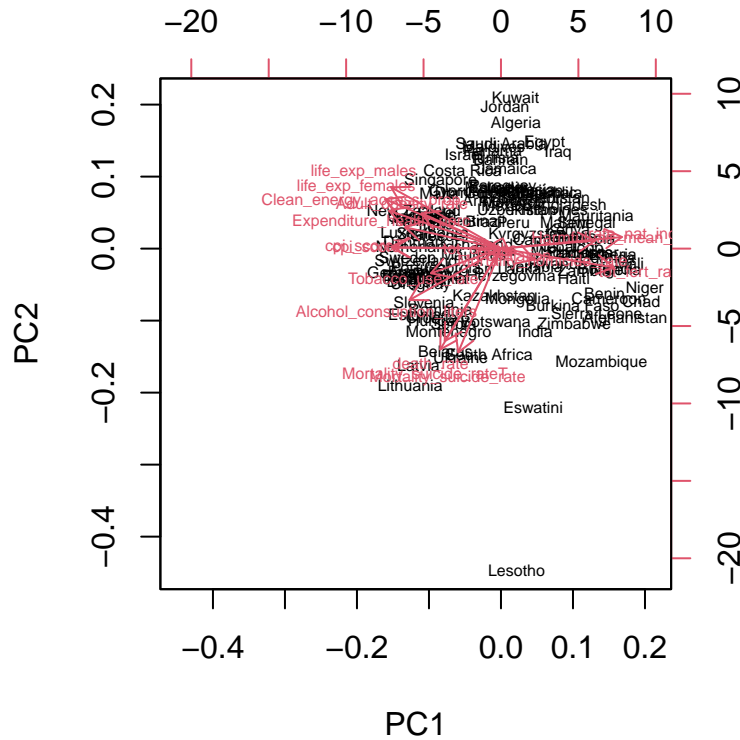
```
ggpairs(df_merge[,c(24,3,25,6,7,10,16, 23)])
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



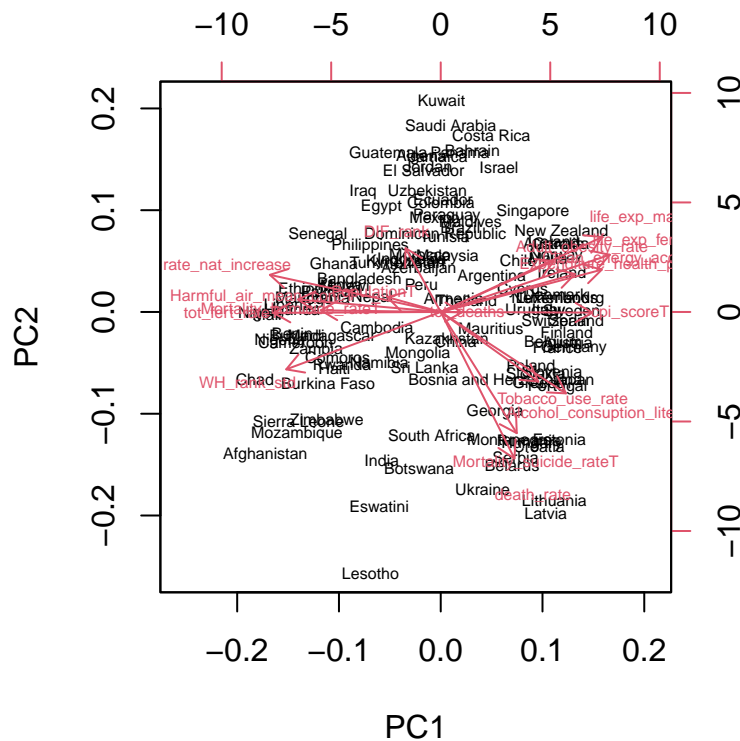
Investigate Groupings by Principal Component Analysis

```
# Non-Transformed Data
pcp = prcomp(df_merge_noNA[, -c(1,16,19,20,21,22,23,24,25,26,27)], center = TRUE, scale = TRUE)
biplot(pcp, cex = 0.5, xlab = df_merge_noNA$country)
```

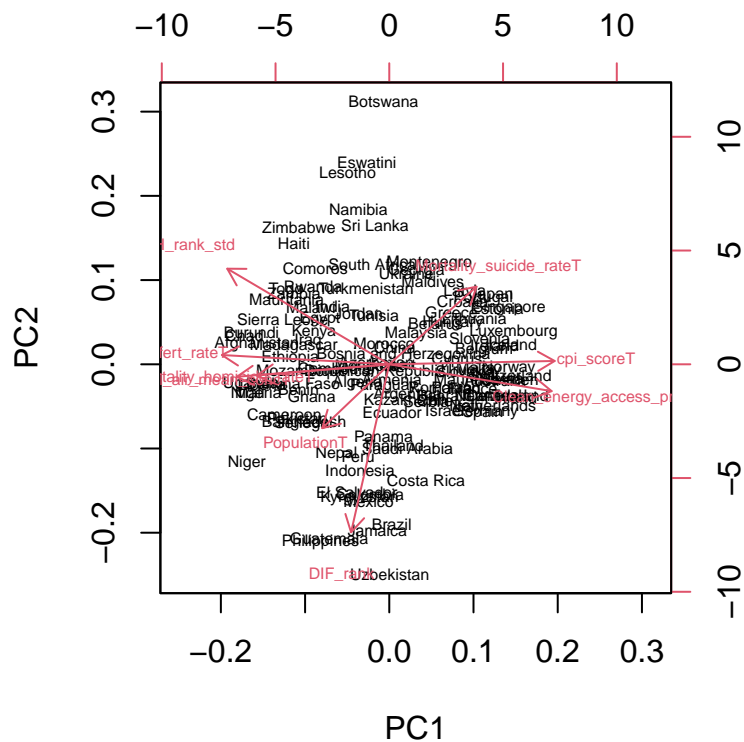
Transformed Data

```
pcpT = prcomp(df_merge_noNA[, -c(1,2,4,9,10,14,18,16,19,20,21)], center = TRUE, scale = TRUE)
biplot(pcpT, cex = 0.5, xlab = df_merge_noNA$country)
```



Final Model (Excluding categorical variable)

```
pcpT = prcomp(df_merge_noNA[, -c(1:7,9:20,21)], center = TRUE, scale = TRUE)
biplot(pcpT, cex = 0.5, xlab = df_merge_noNA$country)
```



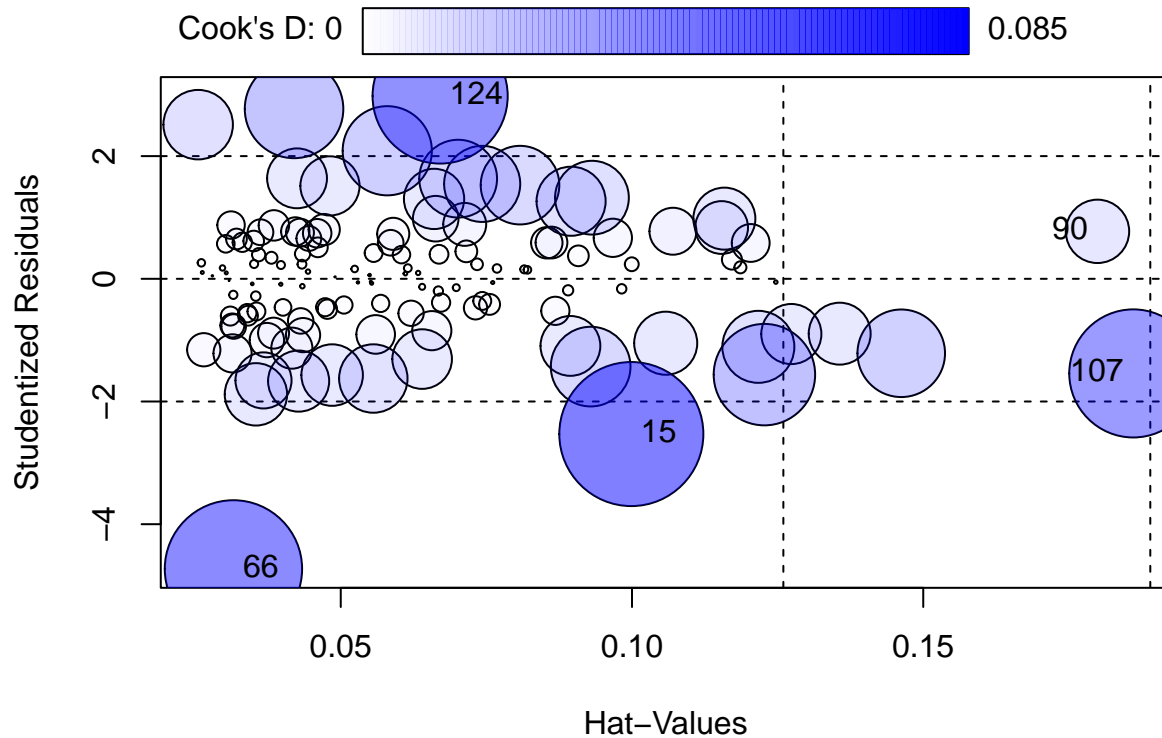
Check Multicollinearity

```
vif(mod_bwAIC)
```

```
##          PopulationT          rate_nat_increase          tot_fert_rateT
##          1.194667          54.416278          32.549900
##          life_exp_females          death_rate Mortality_homicide_rateT
##          5.689041          6.848762          1.742542
##          Univ_hc
##          1.449913
```

Investigate the Countries that are High Leverage Points

```
influencePlot(mod_bwAIC, col=c(1,1))
```



```
##      StudRes      Hat      CookD
## 15  -2.5309970 0.09988576 0.08499729
## 66  -4.7289219 0.03157620 0.07727232
## 90   0.7734527 0.17991352 0.01646078
## 107 -1.5423757 0.18605820 0.06719584
## 124  2.9794075 0.06705869 0.07480577
```

```
"Low Leverage/High Residuals:"
```

```
## [1] "Low Leverage/High Residuals:"
```

```
df_merge[66,1]
```

```
## [1] "Lebanon"
```

```
"High Leverage/Low Residuals:"
```

```
## [1] "High Leverage/Low Residuals:"
```

```
df_merge[90,1]
```

```
## [1] "Niger"
```

```
"High Residual/Low Leverage/High Cooks Distance"
```

```
## [1] "High Residual/Low Leverage/High Cooks Distance"
```

```
df_merge[124,1]
```

```
## [1] "Uzbekistan"
```

```
"High Residual/High Cooks Distance:"
```

```
## [1] "High Residual/High Cooks Distance:"
```

```
df_merge[15,1]
```

```
## [1] "Botswana"
```

```
"High Leverage/High Cooks Distance/Low Residual:"
```

```
## [1] "High Leverage/High Cooks Distance/Low Residual:"
```

```
df_merge[107,1]
```

```
## [1] "Singapore"
```

```
# We see know reason to remove these countries, however we will examine them and also view them in ligh
```

Investigate Subsets of Countries

As the biplot showed, there was clustering of countries based on their economic development status, therefore we are interested in investigating how subsets of countries divided by their economic development status influences the model. We subsetted countries into thirds by their GDP rank, then fitted the final model on each subset.

```
df_merge_first <- subset(df_merge, GDP_rank < 74)
```

```
df_merge_second <- subset(df_merge, GDP_rank < 138 & GDP_rank > 74)
```

```
df_merge_third <- subset(df_merge, GDP_rank > 138)
```

```
"First Subset Model"
```

```
## [1] "First Subset Model"
```

```
mod_first <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_females + death_ra  
summary(mod_first)
```

```
##
```

```
## Call:
```

```
## lm(formula = DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
```

```
##     life_exp_females + death_rate + Mortality_homicide_rateT +
```

```
##     Univ_hc, data = df_merge_first)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -1.90546 -0.11252  0.08376  0.18875  0.50954
```

```
##
```

```
## Coefficients:
```

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

```
## (Intercept)    -2.184875    4.095338  -0.534    0.597
```

```
## PopulationT      0.005154    0.054996   0.094    0.926
```

```
## rate_nat_increase -0.104562    1.185146  -0.088    0.930
```

```
## tot_fert_rateT    1.153312    1.299834   0.887    0.381
```

```
## life_exp_females  0.016144    0.042048   0.384    0.703
```

```
## death_rate       0.012270    0.165180   0.074    0.941
```

```
## Mortality_homicide_rateT -0.075999    0.140899  -0.539    0.593
```

```
## Univ_hc1        -0.099455    0.214106  -0.465    0.645
```

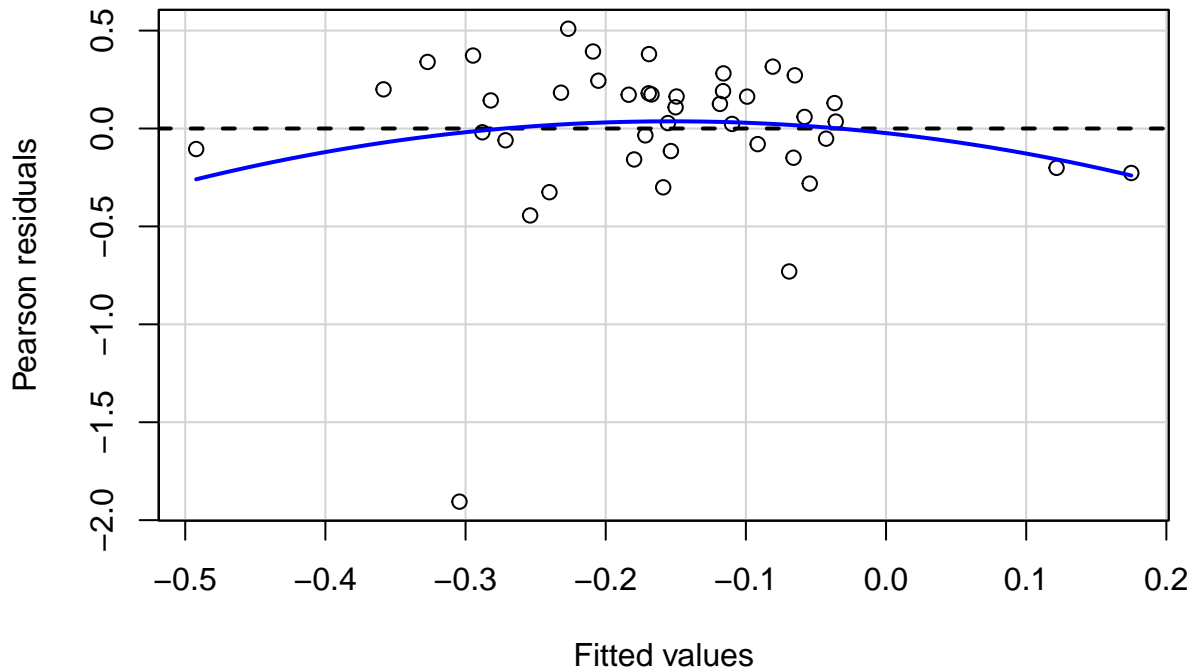
```
##
```

```
## Residual standard error: 0.4293 on 34 degrees of freedom
```

```
## Multiple R-squared:  0.08705,    Adjusted R-squared:  -0.1009
```

```
## F-statistic: 0.4631 on 7 and 34 DF,  p-value: 0.8541
```

```
residualPlot(mod_first)
```



```
"Second Subset Model"
```

```
## [1] "Second Subset Model"
```

```
mod_second <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_females + death_r
summary(mod_second)
```

```
##
```

```
## Call:
```

```
## lm(formula = DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##     life_exp_females + death_rate + Mortality_homicide_rateT +
##     Univ_hc, data = df_merge_second)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.8891 -0.2273 -0.1361  0.2011  1.0950
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -5.18533     4.35620   -1.190    0.242
## PopulationT      0.03170     0.05279    0.601    0.552
## rate_nat_increase -0.26308     1.01664   -0.259    0.797
## tot_fert_rateT   -0.59164     1.41480   -0.418    0.678
## life_exp_females  0.06379     0.03882    1.643    0.110
## death_rate      -0.04726     0.15383   -0.307    0.761
## Mortality_homicide_rateT 0.38440     0.08052    4.774 3.36e-05 ***
## Univ_hc1        -0.23491     0.17179   -1.367    0.180
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

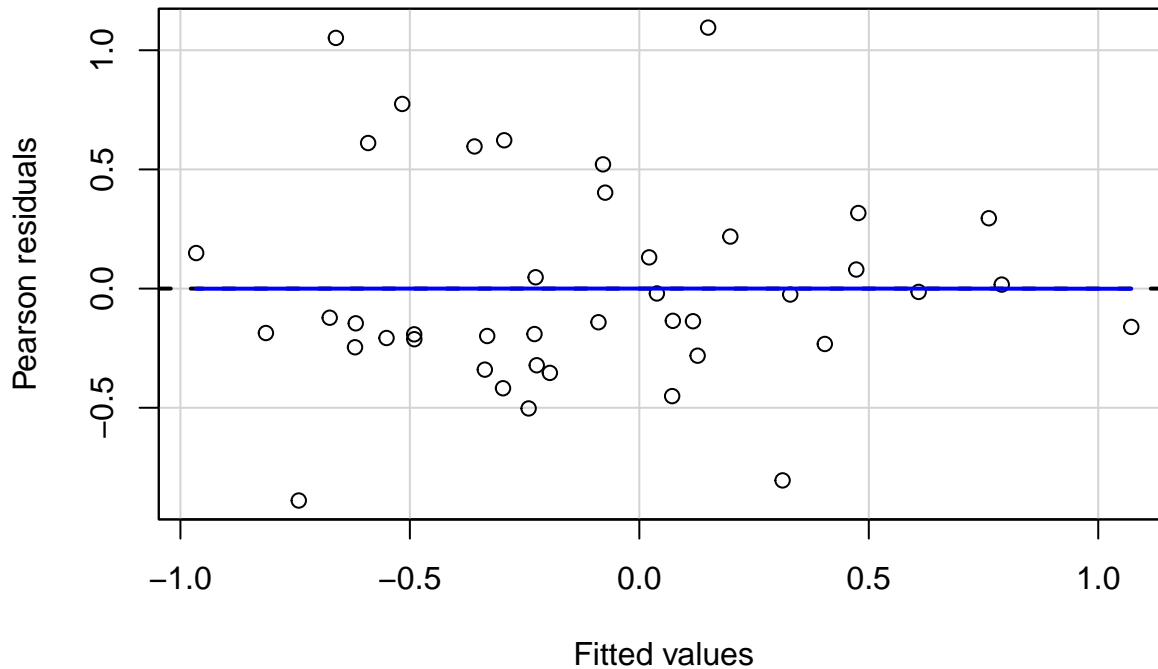
```
##
```

```
## Residual standard error: 0.4759 on 34 degrees of freedom
```

```
## Multiple R-squared:  0.5423, Adjusted R-squared:  0.4481
```

```
## F-statistic: 5.755 on 7 and 34 DF, p-value: 0.0001893
```

```
residualPlot(mod_second)
```



```
"Third Subset Model"
```

```
## [1] "Third Subset Model"
```

```
mod_third <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_females + death_rate + Mortality_homicide_rateT + Univ_hc, data = df_merge_third)
```

```
##
```

```
## Call:
```

```
## lm(formula = DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +  
##      life_exp_females + death_rate + Mortality_homicide_rateT +  
##      Univ_hc, data = df_merge_third)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -0.99014 -0.37787  0.05141  0.35844  1.01142
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)  
## (Intercept)    -3.56093     4.08458  -0.872  0.38943  
## PopulationT     -0.02418     0.07172  -0.337  0.73811  
## rate_nat_increase  1.71256     0.61180   2.799  0.00838 **  
## tot_fert_rateT    -3.54761     1.40059  -2.533  0.01609 *  
## life_exp_females  0.05365     0.04362   1.230  0.22716  
## death_rate       0.19883     0.11848   1.678  0.10247  
## Mortality_homicide_rateT -0.07927     0.11770  -0.673  0.50520  
## Univ_hc1        -0.06718     0.26168  -0.257  0.79894
```

```
## ---
```

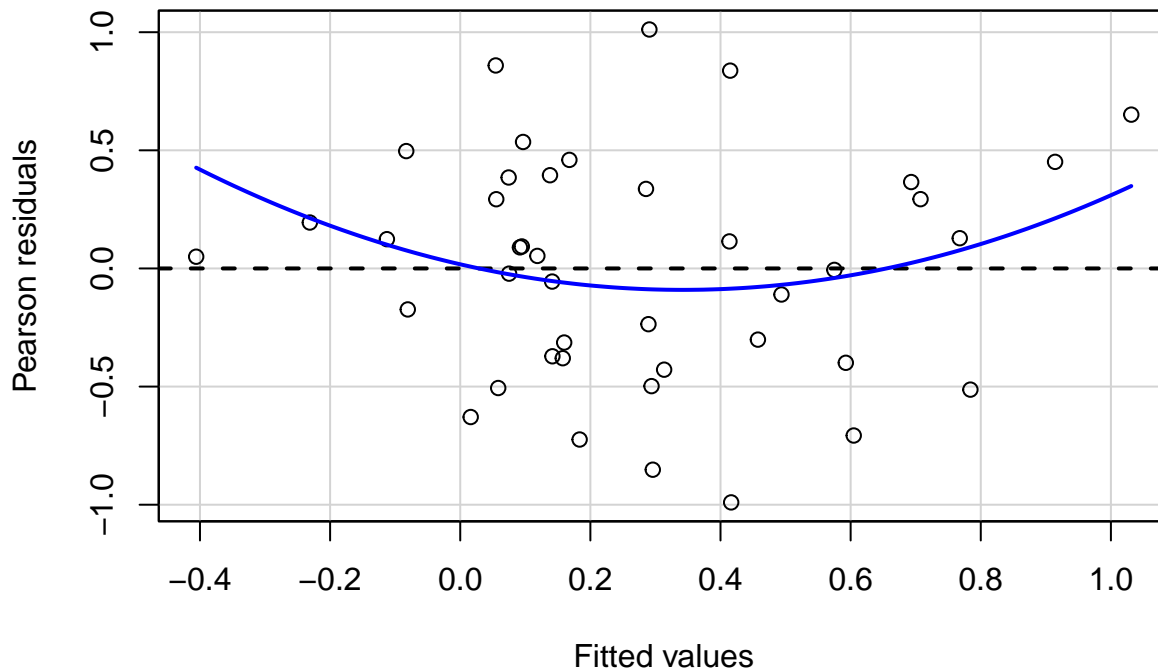
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 0.5278 on 34 degrees of freedom
```

```
## Multiple R-squared:  0.2931, Adjusted R-squared:  0.1475
## F-statistic: 2.014 on 7 and 34 DF,  p-value: 0.08202
```

```
residualPlot(mod_third)
```



The model only fulfills assumptions about residuals for the middle subset of our data.

Remodeling By Subset

```
# Subset 1: Model Selection
```

```
df_merge_firstnoNA <- na.omit(df_merge_first)
```

```
mod_S1 <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_fema
```

```
bw_aicS1 = step(mod_S1, direction = "backward")
```

```
## Start: AIC=-120.71
```

```
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
```

```
## life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
```

```
## Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
```

```
## Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
```

```
## Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT
```

```
##
```

```
##
```

	Df	Sum of Sq	RSS	AIC
## - Mortality_homicide_rateT	1	0.00000	0.79541	-122.71
## - Harmful_air_mean_concT	1	0.00000	0.79541	-122.71
## - death_rate	1	0.00005	0.79546	-122.71
## - tot_fert_rateT	1	0.00014	0.79555	-122.70
## - Alcohol_consuption_liters	1	0.00019	0.79560	-122.70
## - rate_nat_increase	1	0.00119	0.79659	-122.65
## - Expenditure_health_perc	1	0.00498	0.80038	-122.46
## - Tobacco_use_rate	1	0.00769	0.80309	-122.33
## - PopulationT	1	0.02366	0.81907	-121.54
## <none>			0.79541	-120.71

```

## - Adult_obesity_rate      1  0.09800 0.89341 -118.06
## - life_exp_females        1  0.10956 0.90496 -117.55
## - life_exp_males          1  0.14384 0.93924 -116.06
## - Mortality_suicide_rateT 1  0.15883 0.95424 -115.43
## - cpi_scoreT              1  0.18992 0.98533 -114.15
## - tot_deaths              1  0.21389 1.00929 -113.19
## - Univ_hc                 1  0.32561 1.12102 -108.99
## - Clean_energy_access_prop 1  0.33517 1.13057 -108.65
##
## Step: AIC=-122.71
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Adult_obesity_rate + Tobacco_use_rate +
##   Expenditure_health_perc + Mortality_suicide_rateT + Alcohol_consumption_liters +
##   Univ_hc + tot_deaths + cpi_scoreT
##
##           Df Sum of Sq    RSS    AIC
## - Harmful_air_mean_concT  1  0.00000 0.79541 -124.71
## - death_rate              1  0.00007 0.79547 -124.71
## - tot_fert_rateT          1  0.00018 0.79558 -124.70
## - Alcohol_consumption_liters 1  0.00020 0.79560 -124.70
## - rate_nat_increase       1  0.00166 0.79707 -124.63
## - Expenditure_health_perc  1  0.00619 0.80160 -124.40
## - Tobacco_use_rate        1  0.00794 0.80334 -124.31
## - PopulationT             1  0.02368 0.81909 -123.54
## <none>                     0.79541 -122.71
## - life_exp_females        1  0.11452 0.90992 -119.33
## - Adult_obesity_rate      1  0.11573 0.91114 -119.28
## - life_exp_males          1  0.16369 0.95910 -117.23
## - Mortality_suicide_rateT 1  0.16499 0.96039 -117.17
## - cpi_scoreT              1  0.19128 0.98669 -116.09
## - tot_deaths              1  0.21389 1.00929 -115.19
## - Clean_energy_access_prop 1  0.33821 1.13362 -110.54
## - Univ_hc                 1  0.36610 1.16150 -109.57
##
## Step: AIC=-124.71
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Adult_obesity_rate + Tobacco_use_rate + Expenditure_health_perc +
##   Mortality_suicide_rateT + Alcohol_consumption_liters + Univ_hc +
##   tot_deaths + cpi_scoreT
##
##           Df Sum of Sq    RSS    AIC
## - death_rate              1  0.00007 0.79547 -126.71
## - tot_fert_rateT          1  0.00019 0.79560 -126.70
## - Alcohol_consumption_liters 1  0.00024 0.79564 -126.70
## - rate_nat_increase       1  0.00172 0.79713 -126.62
## - Expenditure_health_perc  1  0.00660 0.80200 -126.38
## - Tobacco_use_rate        1  0.00808 0.80348 -126.31
## - PopulationT             1  0.02540 0.82081 -125.45
## <none>                     0.79541 -124.71
## - life_exp_females        1  0.12714 0.92255 -120.78
## - Adult_obesity_rate      1  0.13724 0.93265 -120.34
## - life_exp_males          1  0.16394 0.95934 -119.22

```



```

## - Mortality_suicide_rateT      1    0.17190 0.96731 -118.89
## - tot_deaths                   1    0.22120 1.01660 -116.90
## - cpi_scoreT                   1    0.26676 1.06217 -115.14
## - Clean_energy_access_prop     1    0.35549 1.15090 -111.93
## - Univ_hc                      1    0.36847 1.16388 -111.48
##
## Step:  AIC=-126.71
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + Clean_energy_access_prop +
##   Adult_obesity_rate + Tobacco_use_rate + Expenditure_health_perc +
##   Mortality_suicide_rateT + Alcohol_consuption_liters + Univ_hc +
##   tot_deaths + cpi_scoreT
##
##              Df Sum of Sq      RSS      AIC
## - Alcohol_consuption_liters  1    0.00029 0.79576 -128.69
## - tot_fert_rateT             1    0.00145 0.79692 -128.63
## - Expenditure_health_perc    1    0.00752 0.80299 -128.33
## - Tobacco_use_rate           1    0.00801 0.80348 -128.31
## - PopulationT                1    0.02572 0.82119 -127.44
## - rate_nat_increase          1    0.03219 0.82766 -127.12
## <none>                       0.79547 -126.71
## - life_exp_females           1    0.12977 0.92525 -122.66
## - Adult_obesity_rate         1    0.13816 0.93363 -122.30
## - Mortality_suicide_rateT    1    0.17524 0.97071 -120.74
## - life_exp_males             1    0.19245 0.98793 -120.04
## - tot_deaths                 1    0.22122 1.01669 -118.89
## - cpi_scoreT                 1    0.26684 1.06231 -117.14
## - Clean_energy_access_prop   1    0.36301 1.15849 -113.67
## - Univ_hc                    1    0.40113 1.19661 -112.38
##
## Step:  AIC=-128.69
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + Clean_energy_access_prop +
##   Adult_obesity_rate + Tobacco_use_rate + Expenditure_health_perc +
##   Mortality_suicide_rateT + Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq      RSS      AIC
## - tot_fert_rateT             1    0.00120 0.79697 -130.63
## - Expenditure_health_perc    1    0.00723 0.80300 -130.33
## - Tobacco_use_rate           1    0.00992 0.80568 -130.20
## - PopulationT                1    0.02745 0.82321 -129.34
## - rate_nat_increase          1    0.03940 0.83516 -128.76
## <none>                       0.79576 -128.69
## - life_exp_females           1    0.13565 0.93141 -124.40
## - Adult_obesity_rate         1    0.15788 0.95364 -123.45
## - Mortality_suicide_rateT    1    0.17501 0.97077 -122.74
## - life_exp_males             1    0.19217 0.98794 -122.04
## - tot_deaths                 1    0.22626 1.02203 -120.68
## - cpi_scoreT                 1    0.29211 1.08788 -118.19
## - Clean_energy_access_prop   1    0.40457 1.20034 -114.25
## - Univ_hc                    1    0.40884 1.20460 -114.11
##
## Step:  AIC=-130.63
## DIF_rank ~ PopulationT + rate_nat_increase + life_exp_males +

```

```

##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##      Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq      RSS      AIC
## - Tobacco_use_rate      1    0.00886 0.80582 -132.19
## - Expenditure_health_perc  1    0.00933 0.80630 -132.17
## - PopulationT            1    0.02669 0.82366 -131.31
## <none>                    0.79697 -130.63
## - rate_nat_increase      1    0.05298 0.84995 -130.06
## - life_exp_females        1    0.13614 0.93311 -126.33
## - Adult_obesity_rate      1    0.16459 0.96156 -125.12
## - Mortality_suicide_rateT 1    0.17567 0.97263 -124.67
## - life_exp_males          1    0.19723 0.99419 -123.79
## - tot_deaths              1    0.22526 1.02223 -122.68
## - cpi_scoreT              1    0.29388 1.09085 -120.08
## - Univ_hc                 1    0.41672 1.21369 -115.81
## - Clean_energy_access_prop 1    0.43890 1.23587 -115.08
##
## Step:  AIC=-132.19
## DIF_rank ~ PopulationT + rate_nat_increase + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##              Df Sum of Sq      RSS      AIC
## - Expenditure_health_perc  1    0.01442 0.82025 -133.48
## - PopulationT            1    0.03211 0.83793 -132.63
## <none>                    0.80582 -132.19
## - rate_nat_increase      1    0.04415 0.84997 -132.06
## - life_exp_females        1    0.13111 0.93694 -128.16
## - Adult_obesity_rate      1    0.15609 0.96192 -127.11
## - Mortality_suicide_rateT 1    0.16688 0.97270 -126.66
## - life_exp_males          1    0.19715 1.00298 -125.44
## - tot_deaths              1    0.23674 1.04256 -123.89
## - cpi_scoreT              1    0.28829 1.09412 -121.96
## - Univ_hc                 1    0.43181 1.23764 -117.03
## - Clean_energy_access_prop 1    0.49152 1.29734 -115.14
##
## Step:  AIC=-133.48
## DIF_rank ~ PopulationT + rate_nat_increase + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Mortality_suicide_rateT + Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq      RSS      AIC
## - rate_nat_increase      1    0.03850 0.85875 -133.65
## <none>                    0.82025 -133.48
## - PopulationT            1    0.04709 0.86734 -133.25
## - life_exp_females        1    0.13125 0.95150 -129.54
## - Mortality_suicide_rateT 1    0.15682 0.97707 -128.48
## - Adult_obesity_rate      1    0.15798 0.97823 -128.44
## - life_exp_males          1    0.18707 1.00732 -127.26
## - tot_deaths              1    0.23523 1.05548 -125.39
## - cpi_scoreT              1    0.28583 1.10608 -123.52

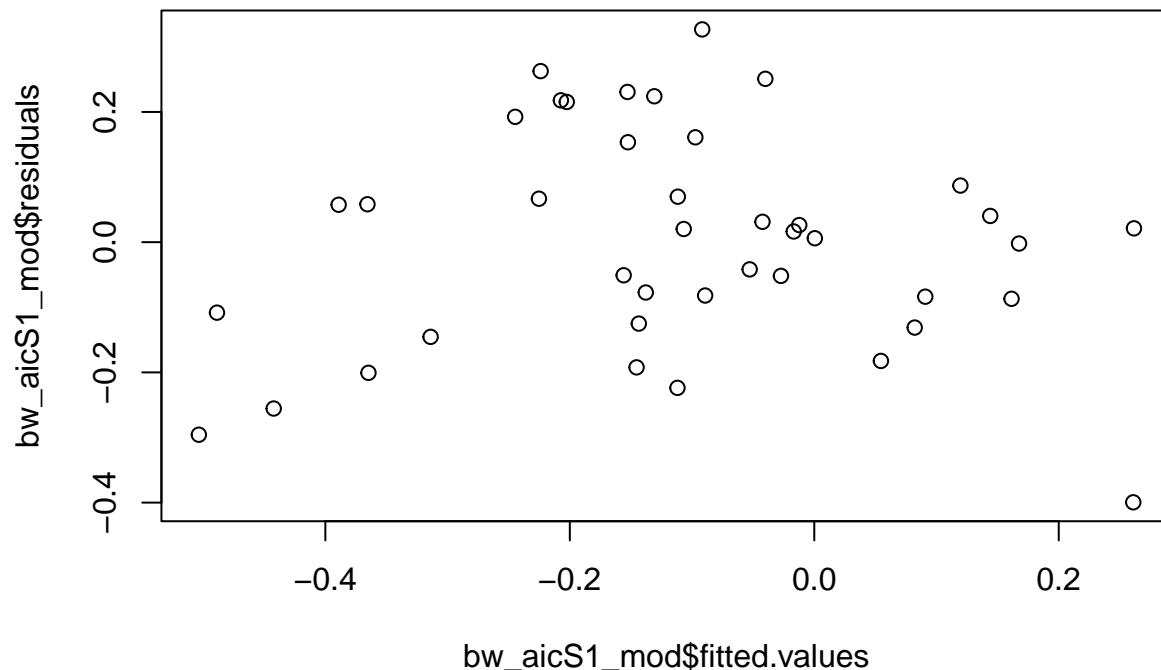
```

```
## - Univ_hc 1 0.42023 1.24048 -118.94
## - Clean_energy_access_prop 1 0.55276 1.37301 -114.88
##
## Step: AIC=-133.65
## DIF_rank ~ PopulationT + life_exp_males + life_exp_females +
## Clean_energy_access_prop + Adult_obesity_rate + Mortality_suicide_rateT +
## Univ_hc + tot_deaths + cpi_scoreT
##
## Df Sum of Sq RSS AIC
## - PopulationT 1 0.04014 0.89889 -133.82
## <none> 0.85875 -133.65
## - Mortality_suicide_rateT 1 0.12076 0.97951 -130.38
## - life_exp_females 1 0.16844 1.02719 -128.48
## - tot_deaths 1 0.19802 1.05677 -127.35
## - Adult_obesity_rate 1 0.24351 1.10226 -125.66
## - life_exp_males 1 0.30380 1.16256 -123.53
## - cpi_scoreT 1 0.34852 1.20727 -122.02
## - Univ_hc 1 0.50333 1.36209 -117.19
## - Clean_energy_access_prop 1 0.62314 1.48189 -113.82
##
## Step: AIC=-133.82
## DIF_rank ~ life_exp_males + life_exp_females + Clean_energy_access_prop +
## Adult_obesity_rate + Mortality_suicide_rateT + Univ_hc +
## tot_deaths + cpi_scoreT
##
## Df Sum of Sq RSS AIC
## <none> 0.89889 -133.82
## - Mortality_suicide_rateT 1 0.10858 1.00747 -131.26
## - life_exp_females 1 0.13578 1.03466 -130.19
## - tot_deaths 1 0.18911 1.08800 -128.18
## - life_exp_males 1 0.26973 1.16862 -125.32
## - Adult_obesity_rate 1 0.29320 1.19209 -124.53
## - cpi_scoreT 1 0.33666 1.23555 -123.09
## - Univ_hc 1 0.58109 1.47998 -115.87
## - Clean_energy_access_prop 1 0.61469 1.51358 -114.98
# ---- Final Model: Mortality_suicide_rateT, life_exp_females, life_exp_males,
# Clean_energy_access_prop, cpi_scoreT, Adult_obesity_rate, Univ_hc
```

```
bw_aicS1_mod <- lm(DIF_rank~life_exp_males + life_exp_females + Clean_energy_access_prop + Adult_obesity_rate + tot_deaths + cpi_scoreT, data = df_merge_first)
summary(bw_aicS1_mod)
```

```
##
## Call:
## lm(formula = DIF_rank ~ life_exp_males + life_exp_females + Clean_energy_access_prop +
## Adult_obesity_rate + Mortality_suicide_rateT + Univ_hc +
## tot_deaths + cpi_scoreT, data = df_merge_first)
##
## Residuals:
## Min 1Q Median 3Q Max
## -0.39955 -0.10827 0.01639 0.08696 0.32678
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.167e+00 1.924e+00 1.127 0.268309
```

```
## life_exp_males      5.844e-02  2.745e-02  2.129 0.041025 *
## life_exp_females   -4.761e-02  3.910e-02 -1.218 0.232272
## Clean_energy_access_prop -5.979e-02  1.450e-02 -4.125 0.000246 ***
## Adult_obesity_rate  2.318e-02  7.617e-03  3.043 0.004649 **
## Mortality_suicide_rateT 1.865e-01  9.376e-02  1.989 0.055357 .
## Univ_hc1           -3.052e-01  8.980e-02 -3.399 0.001828 **
## tot_deaths          1.388e-06  6.719e-07  2.067 0.046952 *
## cpi_scoreT          5.408e-01  2.178e-01  2.484 0.018435 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1874 on 32 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.5598, Adjusted R-squared:  0.4497
## F-statistic: 5.086 on 8 and 32 DF,  p-value: 0.0003893
plot(bw_aicS1_mod$residuals ~ bw_aicS1_mod$fitted.values)
```



```
vif(bw_aicS1_mod)
```

```
##      life_exp_males      life_exp_females Clean_energy_access_prop
##      8.372807      8.828509      1.639454
##      Adult_obesity_rate Mortality_suicide_rateT      Univ_hc
##      2.421033      2.626652      2.117361
##      tot_deaths      cpi_scoreT
##      1.185672      2.975962
```

```
# Subset 2: Model Selection
```

```
df_merge_secondnoNA <- na.omit(df_merge_second)
```

```
mod_S2 <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_fema
```

```
bw_aicS2 = step(mod_S2, direction = "backward")
```

```
## Start:  AIC=-39.13
```

```

## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
##   Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##   Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT
##
##
##      Df Sum of Sq  RSS   AIC
## - Adult_obesity_rate      1  0.00009 4.8571 -41.128
## - rate_nat_increase       1  0.00140 4.8584 -41.118
## - life_exp_males          1  0.00726 4.8642 -41.073
## - death_rate              1  0.00858 4.8656 -41.063
## - PopulationT             1  0.03061 4.8876 -40.896
## - Mortality_suicide_rateT  1  0.05042 4.9074 -40.746
## - Clean_energy_access_prop 1  0.05271 4.9097 -40.729
## - Univ_hc                 1  0.06358 4.9206 -40.647
## - tot_fert_rateT          1  0.08196 4.9389 -40.509
## - Tobacco_use_rate        1  0.08350 4.9405 -40.498
## - tot_deaths              1  0.10664 4.9636 -40.325
## - Harmful_air_mean_concT  1  0.11009 4.9671 -40.299
## - Alcohol_consuption_liters 1  0.25872 5.1157 -39.208
## <none>                     4.8570 -39.129
## - life_exp_females        1  0.27798 5.1350 -39.069
## - Expenditure_health_perc  1  0.44102 5.2980 -37.913
## - cpi_scoreT              1  0.47936 5.3363 -37.646
## - Mortality_homicide_rateT 1  1.81129 6.6683 -29.402
##
## Step:  AIC=-41.13
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##   life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##   Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##   Expenditure_health_perc + Mortality_suicide_rateT + Alcohol_consuption_liters +
##   Univ_hc + tot_deaths + cpi_scoreT
##
##
##      Df Sum of Sq  RSS   AIC
## - rate_nat_increase      1  0.00137 4.8584 -43.117
## - death_rate              1  0.00851 4.8656 -43.063
## - life_exp_males         1  0.00934 4.8664 -43.057
## - PopulationT            1  0.03240 4.8895 -42.882
## - Mortality_suicide_rateT 1  0.05591 4.9130 -42.704
## - Univ_hc                1  0.06778 4.9248 -42.615
## - Tobacco_use_rate       1  0.08686 4.9439 -42.472
## - tot_fert_rateT         1  0.09110 4.9482 -42.440
## - Clean_energy_access_prop 1  0.10082 4.9579 -42.368
## - tot_deaths             1  0.11760 4.9747 -42.243
## - Harmful_air_mean_concT 1  0.13150 4.9886 -42.140
## - Alcohol_consuption_liters 1  0.26255 5.1196 -41.180
## <none>                     4.8571 -41.128
## - life_exp_females        1  0.28618 5.1432 -41.010
## - Expenditure_health_perc  1  0.44999 5.3071 -39.850
## - cpi_scoreT              1  0.48914 5.3462 -39.578
## - Mortality_homicide_rateT 1  2.10173 6.9588 -29.824
##
## Step:  AIC=-43.12
## DIF_rank ~ PopulationT + tot_fert_rateT + life_exp_males + life_exp_females +

```

```

##      death_rate + Clean_energy_access_prop + Harmful_air_mean_concT +
##      Mortality_homicide_rateT + Tobacco_use_rate + Expenditure_health_perc +
##      Mortality_suicide_rateT + Alcohol_consuption_liters + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##
##              Df Sum of Sq    RSS    AIC
## - life_exp_males      1   0.00906 4.8675 -45.049
## - PopulationT         1   0.03105 4.8895 -44.882
## - Mortality_suicide_rateT 1   0.05521 4.9136 -44.699
## - Univ_hc             1   0.06641 4.9248 -44.615
## - Tobacco_use_rate     1   0.09534 4.9538 -44.398
## - Clean_energy_access_prop 1   0.09946 4.9579 -44.368
## - tot_deaths          1   0.11624 4.9747 -44.243
## - Harmful_air_mean_concT 1   0.13338 4.9918 -44.115
## - death_rate          1   0.15952 5.0180 -43.922
## - Alcohol_consuption_liters 1   0.26137 5.1198 -43.179
## <none>                  4.8584 -43.117
## - life_exp_females     1   0.47468 5.3331 -41.668
## - cpi_scoreT           1   0.48779 5.3462 -41.578
## - Expenditure_health_perc 1   0.50268 5.3611 -41.475
## - tot_fert_rateT       1   0.88583 5.7443 -38.920
## - Mortality_homicide_rateT 1   2.13572 6.9942 -31.636
##
## Step:  AIC=-45.05
## DIF_rank ~ PopulationT + tot_fert_rateT + life_exp_females +
##      death_rate + Clean_energy_access_prop + Harmful_air_mean_concT +
##      Mortality_homicide_rateT + Tobacco_use_rate + Expenditure_health_perc +
##      Mortality_suicide_rateT + Alcohol_consuption_liters + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##
##              Df Sum of Sq    RSS    AIC
## - PopulationT         1   0.03630 4.9038 -46.774
## - Mortality_suicide_rateT 1   0.04616 4.9137 -46.699
## - Univ_hc             1   0.07594 4.9434 -46.476
## - Tobacco_use_rate     1   0.08730 4.9548 -46.391
## - Clean_energy_access_prop 1   0.09176 4.9593 -46.358
## - tot_deaths          1   0.14891 5.0164 -45.934
## - Harmful_air_mean_concT 1   0.14942 5.0169 -45.930
## - death_rate          1   0.23558 5.1031 -45.300
## <none>                  4.8675 -45.049
## - Alcohol_consuption_liters 1   0.29622 5.1637 -44.863
## - Expenditure_health_perc 1   0.49637 5.3639 -43.456
## - cpi_scoreT           1   0.50165 5.3691 -43.419
## - tot_fert_rateT       1   0.88299 5.7505 -40.880
## - life_exp_females     1   1.34889 6.2164 -37.998
## - Mortality_homicide_rateT 1   2.13796 7.0054 -33.576
##
## Step:  AIC=-46.77
## DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Alcohol_consuption_liters +
##      Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC

```

```

## - Mortality_suicide_rateT      1    0.06483 4.9686 -48.288
## - Tobacco_use_rate             1    0.08277 4.9866 -48.154
## - Clean_energy_access_prop     1    0.11187 5.0157 -47.939
## - tot_deaths                   1    0.11368 5.0175 -47.926
## - Harmful_air_mean_concT       1    0.11535 5.0192 -47.913
## - Univ_hc                      1    0.15184 5.0556 -47.645
## - death_rate                   1    0.22424 5.1280 -47.119
## - Alcohol_consuption_liters    1    0.27137 5.1752 -46.781
## <none>                          4.9038 -46.774
## - cpi_scoreT                   1    0.46535 5.3692 -45.419
## - Expenditure_health_perc      1    0.52071 5.4245 -45.040
## - tot_fert_rateT               1    0.93958 5.8434 -42.288
## - life_exp_females             1    1.33404 6.2378 -39.870
## - Mortality_homicide_rateT     1    2.10417 7.0080 -35.563
##
## Step:  AIC=-48.29
## Df_rank ~ tot_fert_rateT + life_exp_females + death_rate + Clean_energy_access_prop +
##      Harmful_air_mean_concT + Mortality_homicide_rateT + Tobacco_use_rate +
##      Expenditure_health_perc + Alcohol_consuption_liters + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - Clean_energy_access_prop  1    0.06440 5.0330 -49.811
## - Tobacco_use_rate          1    0.08410 5.0527 -49.667
## - Harmful_air_mean_concT    1    0.14603 5.1147 -49.216
## - Univ_hc                   1    0.16319 5.1318 -49.092
## - tot_deaths                1    0.17455 5.1432 -49.010
## - death_rate                1    0.25280 5.2214 -48.451
## <none>                       4.9686 -48.288
## - cpi_scoreT                1    0.46673 5.4354 -46.966
## - Alcohol_consuption_liters  1    0.67026 5.6389 -45.606
## - Expenditure_health_perc   1    0.72828 5.6969 -45.227
## - tot_fert_rateT            1    0.88543 5.8541 -44.220
## - life_exp_females          1    1.52456 6.4932 -40.386
## - Mortality_homicide_rateT  1    2.05676 7.0254 -37.471
##
## Step:  AIC=-49.81
## Df_rank ~ tot_fert_rateT + life_exp_females + death_rate + Harmful_air_mean_concT +
##      Mortality_homicide_rateT + Tobacco_use_rate + Expenditure_health_perc +
##      Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - Tobacco_use_rate          1    0.10064 5.1337 -51.079
## - Harmful_air_mean_concT    1    0.15398 5.1870 -50.696
## - tot_deaths                1    0.16108 5.1941 -50.646
## - Univ_hc                   1    0.18929 5.2223 -50.445
## - death_rate                1    0.23086 5.2639 -50.152
## <none>                       5.0330 -49.811
## - cpi_scoreT                1    0.47177 5.5048 -48.496
## - Alcohol_consuption_liters  1    0.62621 5.6592 -47.472
## - Expenditure_health_perc   1    0.77726 5.8103 -46.498
## - tot_fert_rateT            1    0.91819 5.9512 -45.611
## - life_exp_females          1    1.46370 6.4967 -42.366
## - Mortality_homicide_rateT  1    2.04271 7.0757 -39.207

```

```

##
## Step: AIC=-51.08
## DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate + Harmful_air_mean_concT +
## Mortality_homicide_rateT + Expenditure_health_perc + Alcohol_consuption_liters +
## Univ_hc + tot_deaths + cpi_scoreT
##
##          Df Sum of Sq    RSS    AIC
## - tot_deaths      1    0.19401  5.3277 -51.706
## - Harmful_air_mean_concT      1    0.20254  5.3362 -51.647
## - Univ_hc          1    0.20776  5.3414 -51.611
## <none>                        5.1337 -51.079
## - death_rate      1    0.33883  5.4725 -50.714
## - cpi_scoreT      1    0.38310  5.5168 -50.416
## - Alcohol_consuption_liters      1    0.63797  5.7716 -48.745
## - Expenditure_health_perc      1    0.72715  5.8608 -48.177
## - tot_fert_rateT      1    0.83402  5.9677 -47.509
## - life_exp_females      1    1.60672  6.7404 -43.004
## - Mortality_homicide_rateT      1    2.28507  7.4187 -39.456
##
## Step: AIC=-51.71
## DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate + Harmful_air_mean_concT +
## Mortality_homicide_rateT + Expenditure_health_perc + Alcohol_consuption_liters +
## Univ_hc + cpi_scoreT
##
##          Df Sum of Sq    RSS    AIC
## - Univ_hc          1    0.0647  5.3924 -53.259
## - Harmful_air_mean_concT      1    0.1306  5.4583 -52.810
## <none>                        5.3277 -51.706
## - death_rate      1    0.3213  5.6490 -51.539
## - Alcohol_consuption_liters      1    0.5366  5.8642 -50.156
## - Expenditure_health_perc      1    0.5424  5.8701 -50.119
## - cpi_scoreT      1    0.5675  5.8951 -49.961
## - tot_fert_rateT      1    0.8672  6.1949 -48.126
## - life_exp_females      1    1.9736  7.3013 -42.046
## - Mortality_homicide_rateT      1    4.3595  9.6872 -31.584
##
## Step: AIC=-53.26
## DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate + Harmful_air_mean_concT +
## Mortality_homicide_rateT + Expenditure_health_perc + Alcohol_consuption_liters +
## cpi_scoreT
##
##          Df Sum of Sq    RSS    AIC
## - Harmful_air_mean_concT      1    0.1259  5.5183 -54.405
## <none>                        5.3924 -53.259
## - death_rate      1    0.3389  5.7313 -53.004
## - Alcohol_consuption_liters      1    0.5000  5.8924 -51.978
## - Expenditure_health_perc      1    0.5006  5.8930 -51.974
## - tot_fert_rateT      1    0.8039  6.1963 -50.118
## - cpi_scoreT      1    0.8233  6.2157 -50.002
## - life_exp_females      1    2.1254  7.5178 -42.965
## - Mortality_homicide_rateT      1    4.4601  9.8525 -32.958
##
## Step: AIC=-54.41
## DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate + Mortality_homicide_rateT +

```



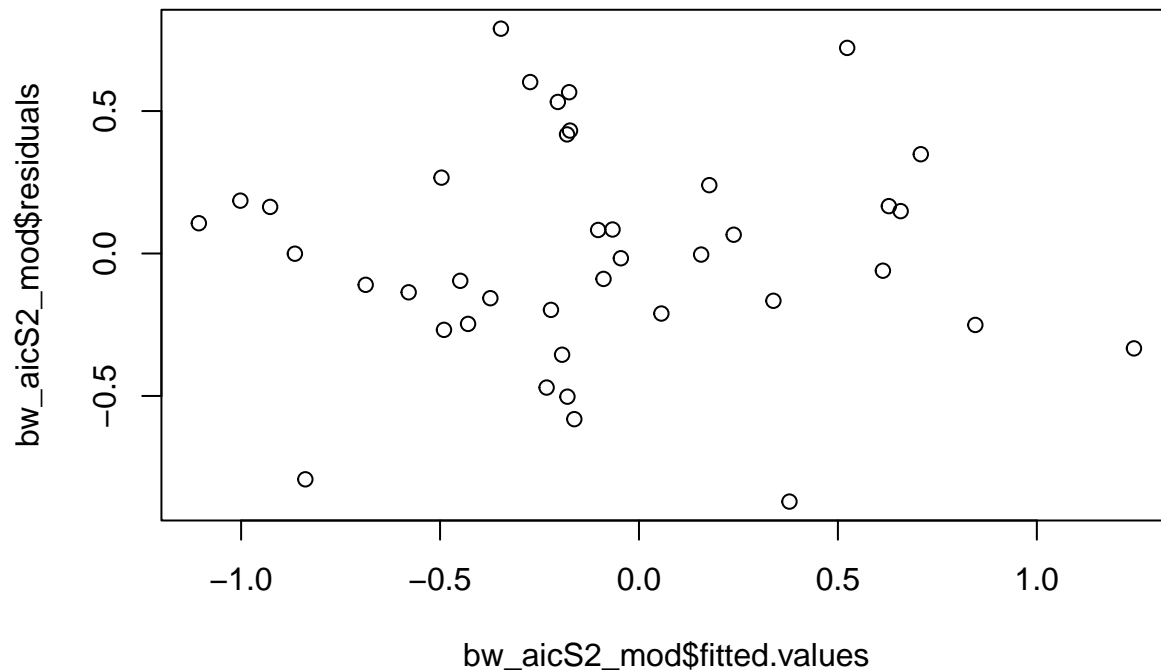
```

##      Expenditure_health_perc + Alcohol_consuption_liters + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## <none>                        5.5183 -54.405
## - death_rate                1    0.3718 5.8901 -53.993
## - Expenditure_health_perc    1    0.4148 5.9331 -53.724
## - Alcohol_consuption_liters  1    0.5218 6.0401 -53.062
## - tot_fert_rateT            1    0.6917 6.2100 -52.036
## - cpi_scoreT                1    0.9142 6.4325 -50.733
## - life_exp_females          1    2.1359 7.6542 -44.300
## - Mortality_homicide_rateT   1    4.3359 9.8542 -34.952
# ---- Final Model: Mortality_suicide_rateT, life_exp_females, death_rate,
# tot_fert_rateT, Clean_energy_access_prop, cpi_scoreT, Mortality_homicide_rateT

bw_aicS2_mod <- lm(DIF_rank~tot_fert_rateT + life_exp_females + death_rate + Mortality_homicide_rateT +
summary(bw_aicS2_mod)

##
## Call:
## lm(formula = DIF_rank ~ tot_fert_rateT + life_exp_females + death_rate +
##      Mortality_homicide_rateT + Expenditure_health_perc + Alcohol_consuption_liters +
##      cpi_scoreT, data = df_merge_second)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87074 -0.21981 -0.01026  0.19907  0.78905
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -3.97294    2.26708  -1.752  0.08928 .
## tot_fert_rateT    -1.04946    0.50692  -2.070  0.04658 *
## life_exp_females     0.07486    0.02150   3.482  0.00146 **
## death_rate        0.05868    0.04552   1.289  0.20657
## Mortality_homicide_rateT  0.40644    0.07555   5.380 6.58e-06 ***
## Expenditure_health_perc  0.02730    0.01952   1.398  0.17160
## Alcohol_consuption_liters -0.05757    0.03202  -1.798  0.08156 .
## cpi_scoreT        -0.67177    0.29525  -2.275  0.02973 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4212 on 32 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.6607, Adjusted R-squared:  0.5865
## F-statistic: 8.902 on 7 and 32 DF,  p-value: 4.705e-06
plot(bw_aicS2_mod$residuals ~ bw_aicS2_mod$fitted.values)

```



```
vif(bw_aicS2_mod)
```

```
##          tot_fert_rateT          life_exp_females          death_rate
##          2.897843          2.038561          3.730778
## Mortality_homicide_rateT Expenditure_health_perc Alcohol_consuption_liters
##          1.710491          1.492865          2.521421
##          cpi_scoreT
##          1.212466
```

```
# Subset 3: Model Selection
```

```
df_merge_thirdnoNA <- na.omit(df_merge_third)
```

```
mod_S3 <- lm(DIF_rank~PopulationT + rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_fema
```

```
bw_aicS3 = step(mod_S3, direction = "backward")
```

```
## Start: AIC=-46.08
```

```
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
## life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
## Harmful_air_mean_concT + Mortality_homicide_rateT + Adult_obesity_rate +
## Tobacco_use_rate + Expenditure_health_perc + Mortality_suicide_rateT +
## Alcohol_consuption_liters + Univ_hc + tot_deaths + cpi_scoreT
##
```

	Df	Sum of Sq	RSS	AIC
## - Harmful_air_mean_concT	1	0.00007	4.0247	-48.083
## - Alcohol_consuption_liters	1	0.00514	4.0298	-48.036
## - PopulationT	1	0.01227	4.0369	-47.971
## - Tobacco_use_rate	1	0.03326	4.0579	-47.779
## - cpi_scoreT	1	0.09595	4.1206	-47.212
## - death_rate	1	0.11101	4.1357	-47.077
## - Univ_hc	1	0.13267	4.1573	-46.884
## - Mortality_homicide_rateT	1	0.18971	4.2144	-46.379
## <none>			4.0247	-46.084
## - Mortality_suicide_rateT	1	0.29213	4.3168	-45.491

```

## - Expenditure_health_perc      1    0.29438 4.3190 -45.472
## - Adult_obesity_rate           1    0.37006 4.3947 -44.829
## - Clean_energy_access_prop     1    0.55398 4.5786 -43.312
## - life_exp_males               1    0.63791 4.6626 -42.640
## - tot_fert_rateT              1    0.68973 4.7144 -42.231
## - life_exp_females            1    1.02324 5.0479 -39.702
## - rate_nat_increase           1    1.07298 5.0976 -39.339
## - tot_deaths                  1    1.13834 5.1630 -38.868
##
## Step:  AIC=-48.08
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Adult_obesity_rate + Tobacco_use_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Alcohol_consumption_liters +
##      Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - Alcohol_consumption_liters  1    0.00537 4.0301 -50.034
## - PopulationT                 1    0.01232 4.0370 -49.970
## - Tobacco_use_rate            1    0.03475 4.0595 -49.765
## - cpi_scoreT                  1    0.09939 4.1241 -49.180
## - death_rate                  1    0.12934 4.1541 -48.913
## - Univ_hc                     1    0.14133 4.1661 -48.806
## - Mortality_homicide_rateT    1    0.19400 4.2187 -48.341
## <none>                        4.0247 -48.083
## - Mortality_suicide_rateT    1    0.29933 4.3241 -47.429
## - Expenditure_health_perc     1    0.35917 4.3839 -46.920
## - Adult_obesity_rate          1    0.43791 4.4626 -46.262
## - Clean_energy_access_prop    1    0.61488 4.6396 -44.823
## - tot_fert_rateT              1    0.72187 4.7466 -43.979
## - life_exp_males              1    0.72282 4.7476 -43.972
## - life_exp_females            1    1.04844 5.0732 -41.517
## - tot_deaths                  1    1.14984 5.1746 -40.785
## - rate_nat_increase           1    1.16767 5.1924 -40.658
##
## Step:  AIC=-50.03
## DIF_rank ~ PopulationT + rate_nat_increase + tot_fert_rateT +
##      life_exp_males + life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Adult_obesity_rate + Tobacco_use_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - PopulationT                 1    0.01238 4.0425 -51.920
## - Tobacco_use_rate            1    0.03378 4.0639 -51.725
## - cpi_scoreT                  1    0.10636 4.1365 -51.070
## - Univ_hc                     1    0.13717 4.1673 -50.795
## - death_rate                  1    0.16086 4.1910 -50.586
## - Mortality_homicide_rateT    1    0.21262 4.2427 -50.131
## <none>                        4.0301 -50.034
## - Mortality_suicide_rateT    1    0.29479 4.3249 -49.422
## - Expenditure_health_perc     1    0.39801 4.4281 -48.549
## - Adult_obesity_rate          1    0.43279 4.4629 -48.260
## - Clean_energy_access_prop    1    0.61045 4.6405 -46.815

```

```

## - tot_fert_rateT      1    0.75487 4.7850 -45.681
## - life_exp_males      1    0.77609 4.8062 -45.518
## - life_exp_females    1    1.04323 5.0733 -43.516
## - rate_nat_increase   1    1.23361 5.2637 -42.153
## - tot_deaths          1    1.25948 5.2896 -41.972
##
## Step:  AIC=-51.92
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Adult_obesity_rate + Tobacco_use_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Univ_hc +
##      tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - Tobacco_use_rate      1    0.02767 4.0701 -53.668
## - cpi_scoreT             1    0.09827 4.1408 -53.032
## - Univ_hc                1    0.17728 4.2198 -52.332
## - death_rate             1    0.17855 4.2210 -52.321
## - Mortality_homicide_rateT 1    0.21878 4.2613 -51.970
## <none>                    4.0425 -51.920
## - Mortality_suicide_rateT 1    0.28643 4.3289 -51.387
## - Expenditure_health_perc 1    0.45574 4.4982 -49.968
## - Adult_obesity_rate     1    0.54914 4.5916 -49.207
## - Clean_energy_access_prop 1    0.66904 4.7115 -48.254
## - tot_fert_rateT         1    0.76209 4.8046 -47.530
## - life_exp_males         1    0.77270 4.8152 -47.448
## - life_exp_females       1    1.03097 5.0735 -45.515
## - rate_nat_increase      1    1.25839 5.3009 -43.893
## - tot_deaths             1    1.27148 5.3140 -43.801
##
## Step:  AIC=-53.67
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Adult_obesity_rate + Expenditure_health_perc +
##      Mortality_suicide_rateT + Univ_hc + tot_deaths + cpi_scoreT
##
##              Df Sum of Sq    RSS    AIC
## - cpi_scoreT             1    0.07941 4.1496 -54.953
## - Univ_hc                1    0.18273 4.2529 -54.043
## - Mortality_homicide_rateT 1    0.20710 4.2772 -53.831
## <none>                    4.0701 -53.668
## - death_rate             1    0.25266 4.3228 -53.439
## - Mortality_suicide_rateT 1    0.29383 4.3640 -53.089
## - Expenditure_health_perc 1    0.43597 4.5061 -51.903
## - Clean_energy_access_prop 1    0.74313 4.8133 -49.463
## - life_exp_males         1    0.76367 4.8338 -49.305
## - tot_fert_rateT         1    0.76823 4.8384 -49.271
## - life_exp_females       1    1.00854 5.0787 -47.477
## - Adult_obesity_rate     1    1.03241 5.1026 -47.303
## - rate_nat_increase      1    1.23074 5.3009 -45.893
## - tot_deaths             1    1.48280 5.5530 -44.174
##
## Step:  AIC=-54.95
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +

```

```

##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Mortality_homicide_rateT + Adult_obesity_rate + Expenditure_health_perc +
##      Mortality_suicide_rateT + Univ_hc + tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - Mortality_homicide_rateT  1    0.21120  4.3608 -55.116
## - death_rate                1    0.22556  4.3751 -54.994
## <none>                      4.1496 -54.953
## - Univ_hc                   1    0.32814  4.4777 -54.137
## - Expenditure_health_perc   1    0.38783  4.5374 -53.647
## - Mortality_suicide_rateT  1    0.40273  4.5523 -53.526
## - life_exp_males            1    0.72598  4.8755 -50.987
## - tot_fert_rateT           1    0.81486  4.9644 -50.319
## - Clean_energy_access_prop  1    0.84947  4.9990 -50.062
## - life_exp_females          1    0.94525  5.0948 -49.360
## - Adult_obesity_rate        1    1.15514  5.3047 -47.866
## - rate_nat_increase         1    1.32304  5.4726 -46.713
## - tot_deaths                1    1.45417  5.6037 -45.837
##
## Step:  AIC=-55.12
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + death_rate + Clean_energy_access_prop +
##      Adult_obesity_rate + Expenditure_health_perc + Mortality_suicide_rateT +
##      Univ_hc + tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - death_rate                1    0.18576  4.5465 -55.573
## <none>                      4.3608 -55.116
## - Univ_hc                   1    0.26877  4.6295 -54.903
## - Mortality_suicide_rateT  1    0.29633  4.6571 -54.684
## - Expenditure_health_perc   1    0.32181  4.6826 -54.482
## - life_exp_males            1    0.55242  4.9132 -52.703
## - life_exp_females          1    0.74782  5.1086 -51.260
## - tot_fert_rateT           1    1.22700  5.5878 -47.943
## - Clean_energy_access_prop  1    1.37697  5.7377 -46.963
## - tot_deaths                1    1.39343  5.7542 -46.857
## - Adult_obesity_rate        1    1.67351  6.0343 -45.098
## - rate_nat_increase         1    1.78715  6.1479 -44.408
##
## Step:  AIC=-55.57
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + Univ_hc +
##      tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - Univ_hc                   1    0.17679  4.7233 -56.161
## - Expenditure_health_perc   1    0.23956  4.7861 -55.673
## <none>                      4.5465 -55.573
## - Mortality_suicide_rateT  1    0.29882  4.8453 -55.217
## - life_exp_females          1    0.56207  5.1086 -53.260
## - life_exp_males            1    0.90045  5.4470 -50.887
## - tot_fert_rateT           1    1.22152  5.7680 -48.768
## - tot_deaths                1    1.23720  5.7837 -48.667

```

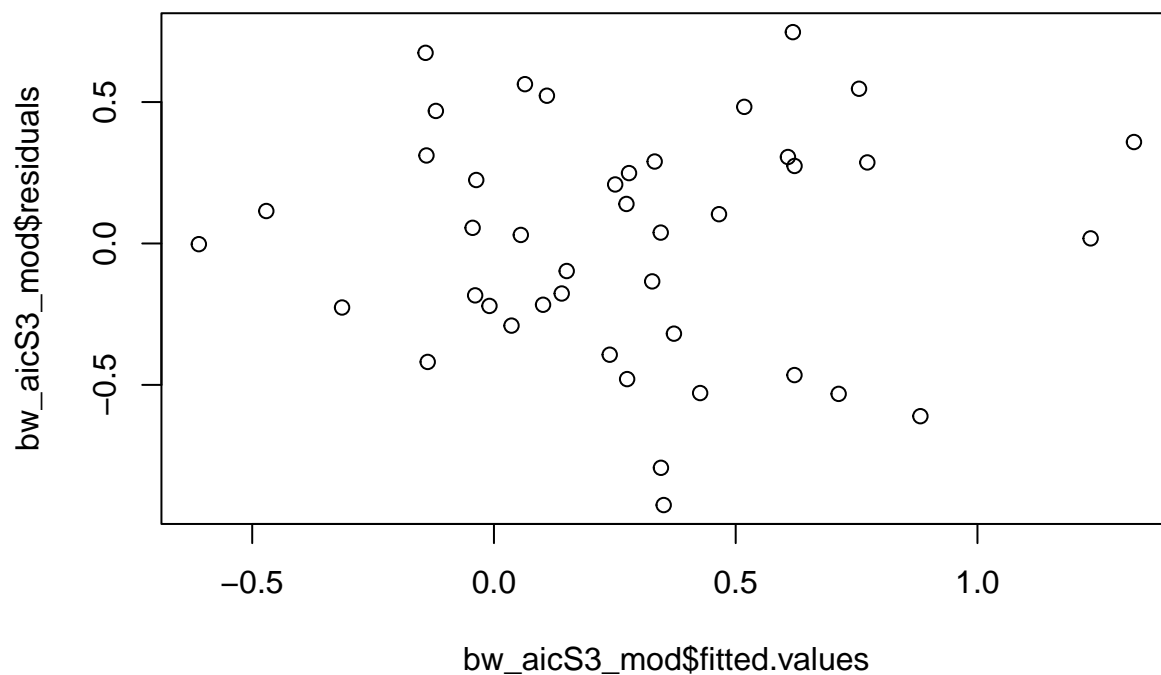
```

## - Adult_obesity_rate      1    1.56510 6.1116 -46.627
## - rate_nat_increase      1    1.69796 6.2445 -45.831
## - Clean_energy_access_prop 1    2.87404 7.4206 -39.447
##
## Step: AIC=-56.16
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Expenditure_health_perc + Mortality_suicide_rateT + tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - Expenditure_health_perc  1    0.13824 4.8615 -57.094
## - Mortality_suicide_rateT  1    0.21601 4.9393 -56.507
## <none>                      4.7233 -56.161
## - life_exp_females         1    0.48483 5.2081 -54.546
## - life_exp_males           1    0.76471 5.4880 -52.609
## - tot_fert_rateT           1    1.05613 5.7794 -50.695
## - tot_deaths               1    1.06088 5.7842 -50.664
## - rate_nat_increase        1    1.52467 6.2480 -47.810
## - Adult_obesity_rate       1    1.55771 6.2810 -47.615
## - Clean_energy_access_prop  1    2.93800 7.6613 -40.265
##
## Step: AIC=-57.09
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      Mortality_suicide_rateT + tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## - Mortality_suicide_rateT  1    0.20612 5.0677 -57.557
## <none>                      4.8615 -57.094
## - life_exp_females         1    0.49157 5.3531 -55.530
## - life_exp_males           1    0.83934 5.7009 -53.201
## - tot_deaths               1    0.94089 5.8024 -52.548
## - tot_fert_rateT           1    1.27079 6.1323 -50.502
## - Adult_obesity_rate       1    1.75914 6.6207 -47.667
## - rate_nat_increase        1    1.78164 6.6432 -47.541
## - Clean_energy_access_prop  1    2.86251 7.7241 -41.963
##
## Step: AIC=-57.56
## DIF_rank ~ rate_nat_increase + tot_fert_rateT + life_exp_males +
##      life_exp_females + Clean_energy_access_prop + Adult_obesity_rate +
##      tot_deaths
##
##              Df Sum of Sq    RSS    AIC
## <none>                      5.0677 -57.557
## - life_exp_females         1    0.42997 5.4976 -56.544
## - tot_deaths               1    0.82446 5.8921 -53.980
## - life_exp_males           1    0.85349 5.9212 -53.798
## - tot_fert_rateT           1    1.34302 6.4107 -50.859
## - Adult_obesity_rate       1    1.58088 6.6485 -49.511
## - rate_nat_increase        1    1.75892 6.8266 -48.533
## - Clean_energy_access_prop  1    2.67314 7.7408 -43.883
# ---- Final Model: Mortality_suicide_rateT, life_exp_females, life_exp_males, tot_deaths,
# rate_nat_increase, tot_fert_rateT, Adult_obesity_rate, Clean_energy_access_prop

```

```
bw_aicS3_mod <- lm(DIF_rank~rate_nat_increase + tot_fert_rateT + life_exp_males + life_exp_females + Cl
summary(bw_aicS3_mod)
```

```
##
## Call:
## lm(formula = DIF_rank ~ rate_nat_increase + tot_fert_rateT +
##     life_exp_males + life_exp_females + Clean_energy_access_prop +
##     Adult_obesity_rate + tot_deaths, data = df_merge_third)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.92482 -0.27439  0.03433  0.28891  0.74720
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.815e+00  1.856e+00   2.056 0.047567 *
## rate_nat_increase  1.886e+00  4.727e-01   3.989 0.000334 ***
## tot_fert_rateT    -4.038e+00  1.080e+00  -3.739 0.000679 ***
## life_exp_males    -1.367e-01  5.904e-02  -2.315 0.026769 *
## life_exp_females   9.013e-02  5.658e-02   1.593 0.120472
## Clean_energy_access_prop  1.497e-02  4.374e-03   3.421 0.001638 **
## Adult_obesity_rate -3.407e-02  1.560e-02  -2.184 0.035961 *
## tot_deaths        -1.768e-06  1.032e-06  -1.713 0.095816 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4453 on 34 degrees of freedom
## Multiple R-squared:  0.4968, Adjusted R-squared:  0.3932
## F-statistic: 4.795 on 7 and 34 DF,  p-value: 0.00078
plot(bw_aicS3_mod$residuals ~ bw_aicS3_mod$fitted.values)
```



```
vif(bw_aicS3_mod)
```

```
##          rate_nat_increase          tot_fert_rateT          life_exp_males
##          28.478025          31.615191          20.731520
##          life_exp_females Clean_energy_access_prop          Adult_obesity_rate
##          22.406557          3.446352          2.094697
##          tot_deaths
##          1.475007
```

```
df_merge_second
```

```
##          country Population rate_nat_increase tot_fert_rate
## 2          Albania    3101621          0.512          1.5459
## 4          Argentina  46621847          0.810          2.1680
## 5          Armenia    2989091          0.126          1.6520
## 8          Azerbaijan  10420515          0.628          1.8587
## 11         Belarus    9383853         -0.394          1.5169
## 14 Bosnia and Herzegovina  3807764         -0.195          1.3718
## 15         Botswana    2417596          1.097          2.3660
## 16         Brazil    218689757          0.654          1.7500
## 17         Bulgaria    6827736         -0.634          1.5073
## 25         China  1413142846          0.187          1.4500
## 26         Colombia    49336454          0.722          1.9434
## 28         Costa Rica    5256612          0.906          1.8565
## 32         Dominican Republic  10790744          1.149          2.1980
## 33         Ecuador    17483326          1.101          2.0180
## 34         Egypt    109546720          1.616          2.7620
## 35         El Salvador    6602370          1.157          2.0360
## 37         Eswatini    1130043          1.330          2.4060
## 41         Gabon    2397368          2.030          3.2600
## 42         Georgia    4936390          0.003          1.7494
## 46         Guatemala    17980803          1.699          2.5700
## 53         Indonesia  279476346          0.828          1.9860
## 54         Iraq    41266109          2.034          3.1740
## 58         Jamaica    2820982          0.833          2.0520
## 60         Jordan    11086716          1.890          2.9140
## 61         Kazakhstan  19543464          0.687          2.0820
## 69         Libya    7252573          1.743          3.0400
## 74         Malaysia    34219975          0.867          1.7403
## 75         Maldives    389568          1.113          1.7040
## 79         Mauritius    1309448          0.087          1.3530
## 80         Mexico    129875529          0.688          1.7348
## 81         Mongolia    3255468          0.904          1.8900
## 82         Montenegro    602445          0.074          1.8073
## 85         Namibia    2777232          1.804          2.9380
## 92         North Macedonia  2133410          0.072          1.5197
## 96         Paraguay    7439863          1.125          1.8849
## 97         Peru    32440172          0.592          2.1760
## 105         Serbia    6693375         -0.625          1.4629
## 110         South Africa  58048332          0.899          2.1660
## 112         Sri Lanka    23326272          0.707          1.9679
## 116         Thailand    69794997          0.218          1.5415
## 119         Turkmenistan    5690818          1.120          2.0280
## 121         Ukraine    43306477         -0.491          1.5664
```


##	life_exp_males	life_exp_females	death_rate	Clean_energy_access_prop
## 2	77.04	82.55	7.36	81
## 4	75.49	81.81	7.28	100
## 5	73.13	79.91	9.54	98
## 8	71.36	77.66	6.92	97
## 11	69.20	80.20	12.81	100
## 14	75.28	81.38	10.26	45
## 15	63.98	68.16	8.98	65
## 16	72.56	79.81	6.90	96
## 17	72.64	79.21	14.31	NA
## 25	75.50	81.20	7.82	79
## 26	71.27	78.69	7.84	93
## 28	77.23	82.65	4.97	96
## 32	71.15	74.60	6.31	92
## 33	75.32	81.32	5.18	94
## 34	73.53	75.98	4.32	100
## 35	72.16	79.27	5.92	92
## 37	58.17	62.33	9.53	55
## 41	68.30	71.81	5.59	88
## 42	73.72	81.96	10.76	89
## 46	71.17	75.30	4.89	50
## 53	71.10	75.68	6.77	84
## 54	71.58	75.42	3.88	99
## 58	74.25	77.85	7.44	83
## 60	74.77	77.84	3.47	100
## 61	67.73	77.56	8.05	93
## 69	75.20	79.76	3.45	NA
## 74	74.77	78.12	5.72	96
## 75	74.84	79.66	4.20	99
## 79	72.32	78.13	8.95	97
## 80	70.29	76.79	7.07	85
## 81	67.49	76.02	6.33	52
## 82	75.57	80.50	10.30	62
## 85	65.00	69.14	6.64	47
## 92	74.99	79.32	9.61	78
## 96	75.97	81.38	4.90	69
## 97	65.38	72.67	11.04	85
## 105	72.46	77.85	15.12	80
## 110	64.26	66.96	9.25	87
## 112	74.83	81.78	6.54	32
## 116	74.92	81.05	7.86	84
## 119	69.10	75.26	5.96	100
## 121	69.10	78.64	13.70	95
##	Harmful_air_mean_conc	Mortality_homicide_rate	Adult_obesity_rate	
## 2	19.3	3.6	21.7	
## 4	12.3	6.1	28.3	
## 5	45.5	3.8	20.2	
## 8	23.2	2.5	19.9	
## 11	18.1	2.7	24.5	
## 14	30.7	1.5	17.9	
## 15	25.0	16.9	18.9	
## 16	11.8	32.6	22.1	
## 17	21.1	1.2	25.0	
## 25	48.8	0.8	6.2	

## 26	19.0	38.3	22.3
## 28	18.6	12.6	25.7
## 32	16.5	17.8	27.6
## 33	19.7	7.0	19.9
## 34	73.0	4.1	32.0
## 35	29.3	85.0	24.6
## 37	19.8	18.5	16.5
## 41	41.5	8.5	15.0
## 42	27.0	2.3	21.7
## 46	30.9	25.1	21.2
## 53	20.7	4.3	6.9
## 54	56.4	14.4	30.4
## 58	15.4	50.3	24.7
## 60	32.8	2.7	35.5
## 61	25.6	5.1	21.0
## 69	46.4	2.1	32.5
## 74	17.1	2.7	15.6
## 75	10.4	1.9	8.6
## 79	9.5	2.3	10.8
## 80	22.8	25.4	28.9
## 81	60.0	6.1	20.6
## 82	19.4	2.8	23.3
## 85	20.5	18.0	17.2
## 92	33.6	1.5	22.4
## 96	12.0	8.0	20.3
## 97	33.3	9.3	19.7
## 105	25.1	1.2	21.5
## 110	29.4	35.9	28.3
## 112	16.8	2.3	5.2
## 116	31.9	4.3	10.0
## 119	33.3	2.8	18.6
## 121	14.9	6.3	24.1
##	Tobacco_use_rate	Expenditure_health_perc	Mortality_suicide_rate
## 2	22.4	NA	4.3
## 4	24.5	15.5	8.4
## 5	25.5	5.7	3.3
## 8	24.0	3.9	4.1
## 11	30.5	11.0	21.2
## 14	35.0	15.4	10.9
## 15	19.4	14.3	16.1
## 16	12.8	10.5	6.9
## 17	39.0	11.6	9.7
## 25	25.6	8.8	8.1
## 26	8.5	16.9	3.9
## 28	8.8	24.1	8.1
## 32	10.6	16.3	4.9
## 33	11.3	13.3	7.6
## 34	24.3	4.7	3.0
## 35	7.9	16.9	6.0
## 37	9.2	10.0	29.4
## 41	NA	9.6	8.4
## 42	31.7	9.4	9.2
## 46	10.9	17.6	5.9
## 53	37.6	8.7	2.4

## 54	18.5		6.0		3.6	
## 58	9.4		13.3		2.4	
## 60	34.8		12.8		1.6	
## 61	23.2		8.2		17.6	
## 69	NA		NA		4.5	
## 74	22.5		8.5		5.7	
## 75	25.2		19.1		2.7	
## 79	20.2		10.2		9.5	
## 80	13.1		10.3		5.3	
## 81	29.4		6.8		17.9	
## 82	31.4		11.3		21.0	
## 85	15.1		10.7		9.7	
## 92	NA		13.6		9.4	
## 96	11.5		14.4		6.0	
## 97	8.1		15.4		2.8	
## 105	39.8		12.0		11.4	
## 110	20.3		15.3		23.5	
## 112	22.0		9.2		14.0	
## 116	22.1		13.9		8.8	
## 119	5.5		8.7		5.7	
## 121	25.8		7.7		21.6	
##	Alcohol_consuption_liters	Univ_hc	tot_deaths	cpi_score	GDP_rank	WH_rank
## 2	6.8	1	3596	36	109	91
## 4	9.5	1	130124	38	84	56
## 5	4.7	0	8710	46	121	84
## 8	1.0	0	10008	23	117	88
## 11	11.0	0	7118	39	103	73
## 14	7.8	0	16234	34	102	63
## 15	6.6	1	2786	60	106	143
## 16	7.3	1	693734	38	101	34
## 17	12.5	1	38106	43	79	86
## 25	6.0	1	48738	45	76	82
## 26	5.5	1	142179	39	112	51
## 28	4.1	1	9085	54	75	16
## 32	6.7	0	4384	32	96	72
## 33	3.3	0	35948	36	114	65
## 34	0.1	1	24802	30	137	129
## 35	4.1	0	4230	33	127	48
## 37	8.8	0	1422	30	133	127
## 41	8.1	0	306	29	98	110
## 42	9.5	1	16903	56	120	106
## 46	1.6	0	19998	24	122	30
## 53	0.2	1	160612	34	131	80
## 54	0.4	0	25374	23	124	109
## 58	4.2	0	3463	44	119	36
## 60	0.5	0	14122	47	132	124
## 61	5.0	0	19058	36	86	44
## 69	0.0	0	6437	17	115	78
## 74	0.9	1	36853	47	83	79
## 75	2.8	1	311	40	85	87
## 79	4.8	1	1040	50	93	49
## 80	5.0	1	331494	31	87	35
## 81	5.9	0	2136	33	128	69
## 82	12.2	0	2791	45	90	71

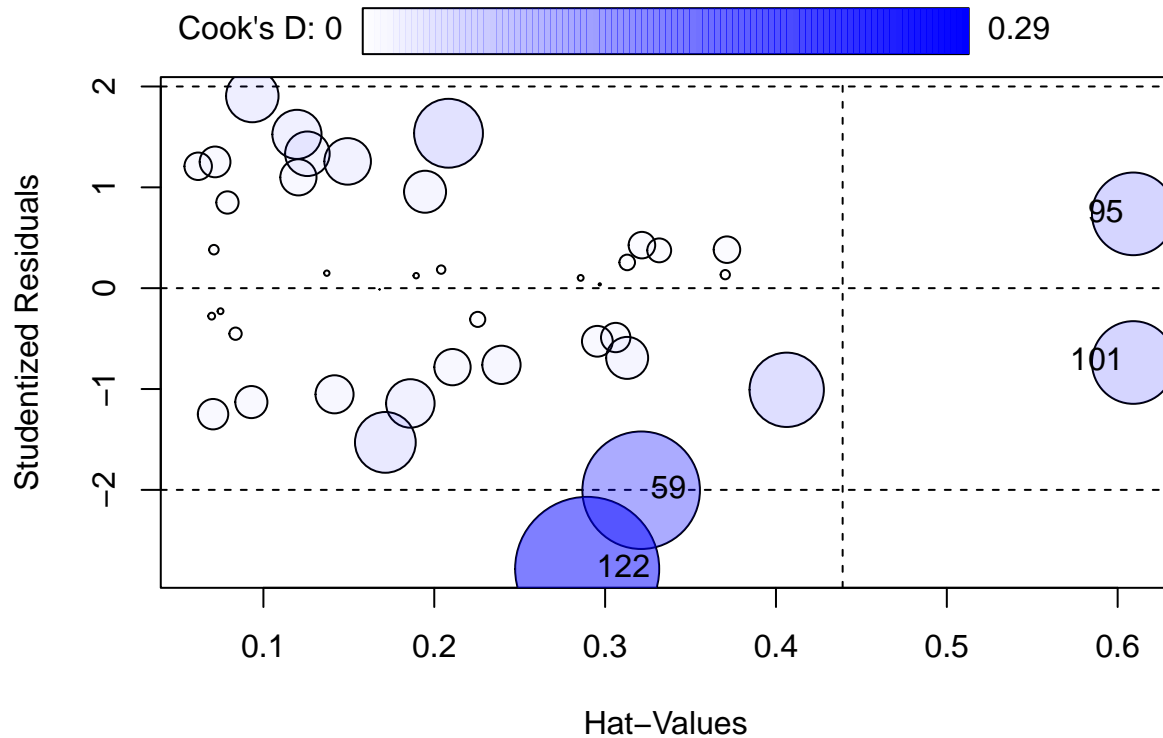
## 85		3.1	0	4082	49	123	122
## 92		6.4	0	9614	40	108	92
## 96		7.0	0	19688	28	113	70
## 97		6.8	1	218178	36	107	62
## 105		8.9	1	17515	36	94	47
## 110		9.5	1	102568	43	105	101
## 112		2.9	1	16817	36	136	126
## 116		8.5	1	33669	36	104	53
## 119		3.1	0	0	19	95	95
## 121		8.3	0	110813	33	126	108
##	GDP_rank_std	WH_rank_std	DIF_rank	PopulationT	tot_fert_rateT		
## 2	0.03294622	0.41251058	-0.37956435	14.94744	0.4356063		
## 4	-0.37888157	-0.39667279	0.01779122	17.65758	0.7738051		
## 5	0.23062356	0.25067390	-0.02005034	14.91048	0.5019867		
## 8	0.16473112	0.34315200	-0.17842089	16.15929	0.6198773		
## 11	-0.06589245	-0.00364087	-0.06225158	16.05450	0.4166688		
## 14	-0.08236556	-0.23483612	0.15247056	15.15255	0.3161237		
## 15	-0.01647311	1.61472587	-1.63119898	14.69828	0.8612008		
## 16	-0.09883867	-0.90530234	0.80646367	19.20316	0.5596158		
## 17	-0.46124713	0.29691295	-0.75816008	15.73650	0.4103200		
## 25	-0.51066646	0.20443485	-0.71510131	21.06908	0.3715636		
## 26	0.08236556	-0.51227042	0.59463597	17.71417	0.6644390		
## 28	-0.52713957	-1.32145378	0.79431421	15.47500	0.6186930		
## 32	-0.18120423	-0.02676039	-0.15444383	16.19420	0.7875479		
## 33	0.11531178	-0.18859707	0.30390885	16.67676	0.7021069		
## 34	0.49419335	1.29105252	-0.79685917	18.51186	1.0159551		
## 35	0.32946223	-0.58162899	0.91109122	15.70294	0.7109871		
## 37	0.42830090	1.24481347	-0.81651257	13.93777	0.8779656		
## 41	-0.14825800	0.85178155	-1.00003955	14.68988	1.1817272		
## 42	0.21415045	0.75930345	-0.54515300	15.41214	0.5592729		
## 46	0.24709667	-0.99778044	1.24487711	16.70482	0.9439059		
## 53	0.39535468	0.15819580	0.23715888	19.44843	0.6861226		
## 54	0.28004290	0.82866202	-0.54861912	17.53555	1.1549926		
## 58	0.19767734	-0.85906329	1.05674063	14.85260	0.7188149		
## 60	0.41182779	1.17545490	-0.76362710	16.22126	1.0695267		
## 61	-0.34593534	-0.67410709	0.32817174	16.78815	0.7333290		
## 69	0.13178489	0.11195675	0.01982814	15.79687	1.1118575		
## 74	-0.39535468	0.13507628	-0.53043096	17.34832	0.5540575		
## 75	-0.36240846	0.32003248	-0.68244093	12.87279	0.5329784		
## 79	-0.23062356	-0.55850947	0.32788590	14.08512	0.3023243		
## 80	-0.32946223	-0.88218281	0.55272058	18.68209	0.5508921		
## 81	0.34593534	-0.09611897	0.44205431	14.99585	0.6365768		
## 82	-0.28004290	-0.04987992	-0.23016298	13.30875	0.5918340		
## 85	0.26356979	1.12921585	-0.86564606	14.83697	1.0777291		
## 92	0.01647311	0.43563010	-0.41915699	14.57323	0.4185129		
## 96	0.09883867	-0.07299944	0.17183811	15.82236	0.6338748		
## 97	0.00000000	-0.25795564	0.25795564	17.29491	0.7774883		
## 105	-0.21415045	-0.60474852	0.39059806	15.71663	0.3804208		
## 110	-0.03294622	0.64370582	-0.67665205	17.87679	0.7728821		
## 112	0.47772024	1.22169394	-0.74397371	16.96509	0.6769670		
## 116	-0.04941933	-0.46603137	0.41661203	18.06107	0.4327560		
## 119	-0.19767734	0.50498868	-0.70266602	15.55436	0.7070501		
## 121	0.31298912	0.80554250	-0.49255338	17.58381	0.4487800		
##	Harmful_air_mean_concT	Mortality_homicide_rateT	Mortality_suicide_rateT				

## 2	2.960105	1.2809338	1.4586150
## 4	2.509599	1.8082888	2.1282317
## 5	3.817712	1.3350011	1.1939225
## 8	3.144152	0.9162907	1.4109870
## 11	2.895912	0.9932518	3.0540012
## 14	3.424263	0.4054651	2.3887628
## 15	3.218876	2.8273136	2.7788193
## 16	2.468100	3.4843123	1.9315214
## 17	3.049273	0.1823216	2.2721259
## 25	3.887730	-0.2231436	2.0918641
## 26	2.944439	3.6454499	1.3609766
## 28	2.923162	2.5336968	2.0918641
## 32	2.803360	2.8791985	1.5892352
## 33	2.980619	1.9459101	2.0281482
## 34	4.290459	1.4109870	1.0986123
## 35	3.377588	4.4426513	1.7917595
## 37	2.985682	2.9177707	3.3809947
## 41	3.725693	2.1400662	2.1282317
## 42	3.295837	0.8329091	2.2192035
## 46	3.430756	3.2228678	1.7749524
## 53	3.030134	1.4586150	0.8754687
## 54	4.032469	2.6672282	1.2809338
## 58	2.734368	3.9180051	0.8754687
## 60	3.490429	0.9932518	0.4700036
## 61	3.242592	1.6292405	2.8678989
## 69	3.837299	0.7419373	1.5040774
## 74	2.839078	0.9932518	1.7404662
## 75	2.341806	0.6418539	0.9932518
## 79	2.251292	0.8329091	2.2512918
## 80	3.126761	3.2347492	1.6677068
## 81	4.094345	1.8082888	2.8848007
## 82	2.965273	1.0296194	3.0445224
## 85	3.020425	2.8903718	2.2721259
## 92	3.514526	0.4054651	2.2407097
## 96	2.484907	2.0794415	1.7917595
## 97	3.505557	2.2300144	1.0296194
## 105	3.222868	0.1823216	2.4336134
## 110	3.380995	3.5807373	3.1570004
## 112	2.821379	0.8329091	2.6390573
## 116	3.462606	1.4586150	2.1747517
## 119	3.505557	1.0296194	1.7404662
## 121	2.701361	1.8405496	3.0726933
##	cpi_scoreT		
## 2	3.583519		
## 4	3.637586		
## 5	3.828641		
## 8	3.135494		
## 11	3.663562		
## 14	3.526361		
## 15	4.094345		
## 16	3.637586		
## 17	3.761200		
## 25	3.806662		
## 26	3.663562		

```
## 28      3.988984
## 32      3.465736
## 33      3.583519
## 34      3.401197
## 35      3.496508
## 37      3.401197
## 41      3.367296
## 42      4.025352
## 46      3.178054
## 53      3.526361
## 54      3.135494
## 58      3.784190
## 60      3.850148
## 61      3.583519
## 69      2.833213
## 74      3.850148
## 75      3.688879
## 79      3.912023
## 80      3.433987
## 81      3.496508
## 82      3.806662
## 85      3.891820
## 92      3.688879
## 96      3.332205
## 97      3.583519
## 105     3.583519
## 110     3.761200
## 112     3.583519
## 116     3.583519
## 119     2.944439
## 121     3.496508
```

```
## Looking at remaining outliers from individual models
```

```
influencePlot(bw_aicS1_mod, col=c(1,1))
```



```
##      StudRes      Hat      CookD
## 59  -2.0035781 0.3211073 0.19280719
## 95   0.7369986 0.6091450 0.09542023
## 101 -0.7369986 0.6091450 0.09542023
## 122 -2.7836429 0.2894979 0.28970627
```

```
"High Residual, high cooks distance"
```

```
## [1] "High Residual, high cooks distance"
```

```
df_merge[122, 1]
```

```
## [1] "United Arab Emirates"
```

```
df_merge[59, 1]
```

```
## [1] "Japan"
```

```
"High Leverage, low residual, low studentized residual"
```

```
## [1] "High Leverage, low residual, low studentized residual"
```

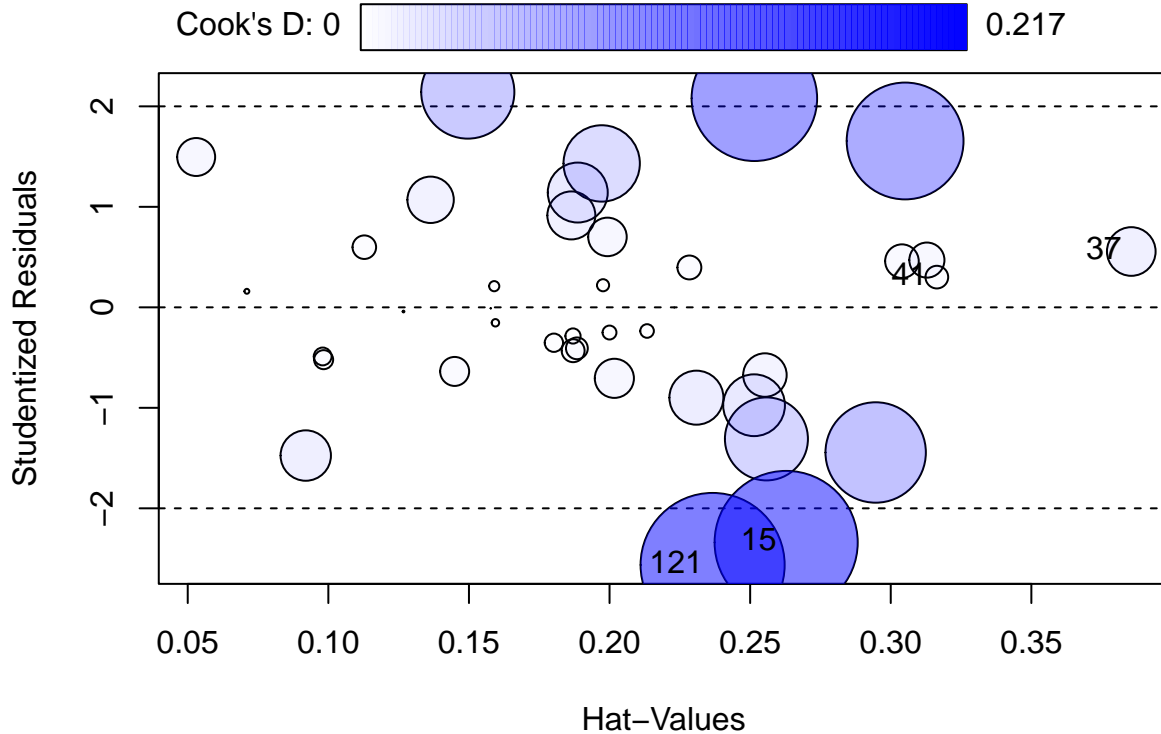
```
df_merge[95, 1]
```

```
## [1] "Panama"
```

```
df_merge[101, 1]
```

```
## [1] "Romania"
```

```
influencePlot(bw_aicS2_mod, col=c(1,1))
```



```
##      StudRes      Hat      CookD
## 15  -2.3395278 0.2628064 0.213990552
## 37   0.5555602 0.3855372 0.024741654
## 41   0.3006099 0.3163881 0.005380854
## 121 -2.5635556 0.2366514 0.216904887
```

```
"High Residual, high cooks distance"
```

```
## [1] "High Residual, high cooks distance"
```

```
df_merge[121, 1]
```

```
## [1] "Ukraine"
```

```
df_merge[15, 1]
```

```
## [1] "Botswana"
```

```
"High Leverage, low residual, low studentized residual"
```

```
## [1] "High Leverage, low residual, low studentized residual"
```

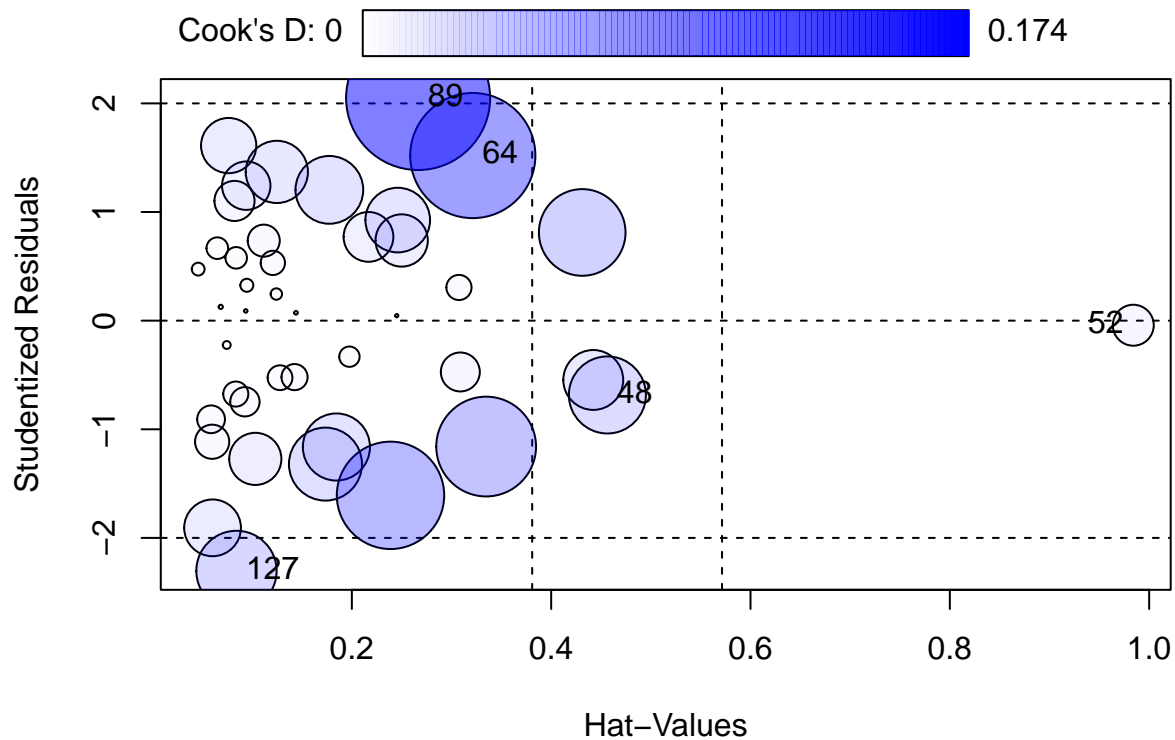
```
df_merge[37, 1]
```

```
## [1] "Eswatini"
```

```
df_merge[41, 1]
```

```
## [1] "Gabon"
```

```
influencePlot(bw_aicS3_mod, col=c(1,1))
```

```
##      StudRes      Hat      CookD
## 48  -0.68397958 0.45641117 0.04988080
## 52  -0.04211378 0.98406973 0.01410921
## 64   1.51970045 0.32136960 0.13163912
## 89   2.04938591 0.26649790 0.17433562
## 127 -2.30370653 0.08425636 0.05417428
```

```
"High Residual, high cooks distance, low leverage"
```

```
## [1] "High Residual, high cooks distance, low leverage"
```

```
df_merge[127, 1]
```

```
## [1] "Zimbabwe"
```

```
df_merge[89, 1]
```

```
## [1] "Nicaragua"
```

```
df_merge[64, 1]
```

```
## [1] "Kyrgyzstan"
```

```
"High Leverage, low residual, low studentized residual"
```

```
## [1] "High Leverage, low residual, low studentized residual"
```

```
df_merge[52, 1]
```

```
## [1] "India"
```

Model Diagnostics

```
par(mfrow=c(2,2))
plot(mod_bwAIC)
```

