Final Project

Sai Akshay Reddy Mavuram, Sai Sreekar Jakku, Ajay Kumar Mashapari

2024-04-24

Importing Packages

```
library(ggplot2)
library(MLmetrics)
```

```
Warning: package 'MLmetrics' was built under R version 4.3.3
```

```

Attaching package: 'MLmetrics'
```

```
The following object is masked from 'package:base':

    Recall
```

```
library(MASS)
```

```
Warning: package 'MASS' was built under R version 4.3.3
```

Data Cleaning and Preparation

Importing Data Set

```
WHO <- read.csv("WHO.csv", sep = ",")
head(WHO)
```

```
      Country Year     Status Life.expectancy Adult.Mortality infant.deaths
1 Afghanistan 2015 Developing            65.0             263            62
2 Afghanistan 2014 Developing            59.9             271            64
3 Afghanistan 2013 Developing            59.9             268            66
4 Afghanistan 2012 Developing            59.5             272            69
5 Afghanistan 2011 Developing            59.2             275            71
6 Afghanistan 2010 Developing            58.8             279            74
  Alcohol percentage.expenditure Hepatitis.B Measles  BMI under.five.deaths
1    0.01              71.279624          65    1154 19.1                83
2    0.01              73.523582          62     492 18.6                86
3    0.01              73.219243          64     430 18.1                89
4    0.01              78.184215          67    2787 17.6                93
5    0.01               7.097109          68    3013 17.2                97
6    0.01              79.679367          66    1989 16.7               102
  Polio Total.expenditure Diphtheria HIV.AIDS       GDP Population
1     6              8.16         65      0.1 584.25921   33736494
2    58              8.18         62      0.1 612.69651     327582
3    62              8.13         64      0.1 631.74498   31731688
```

```
4    67              8.52         67      0.1 669.95900    3696958
5    68              7.87         68      0.1  63.53723    2978599
6    66              9.20         66      0.1 553.32894    2883167
  thinness..1.19.years thinness.5.9.years Income.composition.of.resources
1                 17.2              17.3                           0.479
2                 17.5              17.5                           0.476
3                 17.7              17.7                           0.470
4                 17.9              18.0                           0.463
5                 18.2              18.2                           0.454
6                 18.4              18.4                           0.448
  Schooling
1      10.1
2      10.0
3       9.9
4       9.8
5       9.5
6       9.2
```

```
summary(WHO)
```

```
   Country               Year         Status          Life.expectancy
 Length:2938        Min.   :2000   Length:2938        Min.   :36.30
 Class :character   1st Qu.:2004   Class :character   1st Qu.:63.10
 Mode  :character   Median :2008   Mode  :character   Median :72.10
                    Mean   :2008                      Mean   :69.22
                    3rd Qu.:2012                      3rd Qu.:75.70
                    Max.   :2015                      Max.   :89.00
                                                      NA's   :10
 Adult.Mortality infant.deaths       Alcohol        percentage.expenditure
 Min.   :  1.0   Min.   :   0.0   Min.   : 0.0100   Min.   :    0.000
 1st Qu.: 74.0   1st Qu.:   0.0   1st Qu.: 0.8775   1st Qu.:    4.685
 Median :144.0   Median :   3.0   Median : 3.7550   Median :   64.913
 Mean   :164.8   Mean   :  30.3   Mean   : 4.6029   Mean   :  738.251
 3rd Qu.:228.0   3rd Qu.:  22.0   3rd Qu.: 7.7025   3rd Qu.:  441.534
 Max.   :723.0   Max.   :1800.0   Max.   :17.8700   Max.   :19479.912
 NA's   :10                       NA's   :194
  Hepatitis.B       Measles               BMI        under.five.deaths
 Min.   : 1.00   Min.   :     0.0   Min.   : 1.00   Min.   :   0.00
 1st Qu.:77.00   1st Qu.:     0.0   1st Qu.:19.30   1st Qu.:   0.00
 Median :92.00   Median :    17.0   Median :43.50   Median :   4.00
 Mean   :80.94   Mean   :  2419.6   Mean   :38.32   Mean   :  42.04
 3rd Qu.:97.00   3rd Qu.:   360.2   3rd Qu.:56.20   3rd Qu.:  28.00
 Max.   :99.00   Max.   :212183.0   Max.   :87.30   Max.   :2500.00
 NA's   :553                        NA's   :34
     Polio       Total.expenditure   Diphtheria       HIV.AIDS
 Min.   : 3.00   Min.   : 0.370    Min.   : 2.00   Min.   : 0.100
 1st Qu.:78.00   1st Qu.: 4.260    1st Qu.:78.00   1st Qu.: 0.100
 Median :93.00   Median : 5.755    Median :93.00   Median : 0.100
 Mean   :82.55   Mean   : 5.938    Mean   :82.32   Mean   : 1.742
 3rd Qu.:97.00   3rd Qu.: 7.492    3rd Qu.:97.00   3rd Qu.: 0.800
```

```
 Max.   :99.00   Max.   :17.600    Max.   :99.00   Max.   :50.600
 NA's   :19      NA's   :226       NA's   :19
      GDP              Population        thinness..1.19.years
 Min.   :     1.68   Min.   :3.400e+01   Min.   : 0.10
 1st Qu.:   463.94   1st Qu.:1.958e+05   1st Qu.: 1.60
 Median :  1766.95   Median :1.387e+06   Median : 3.30
 Mean   :  7483.16   Mean   :1.275e+07   Mean   : 4.84
 3rd Qu.:  5910.81   3rd Qu.:7.420e+06   3rd Qu.: 7.20
 Max.   :119172.74   Max.   :1.294e+09   Max.   :27.70
 NA's   :448         NA's   :652         NA's   :34
 thinness.5.9.years Income.composition.of.resources   Schooling
 Min.   : 0.10      Min.   :0.0000                    Min.   : 0.00
 1st Qu.: 1.50      1st Qu.:0.4930                    1st Qu.:10.10
 Median : 3.30      Median :0.6770                    Median :12.30
 Mean   : 4.87      Mean   :0.6276                    Mean   :11.99
 3rd Qu.: 7.20      3rd Qu.:0.7790                    3rd Qu.:14.30
 Max.   :28.60      Max.   :0.9480                    Max.   :20.70
 NA's   :34         NA's   :167                       NA's   :163
```

Checking Data Types

```
str(WHO)
```

```
'data.frame':   2938 obs. of  22 variables:
 $ Country                        : chr  "Afghanistan" "Afghanistan"
"Afghanistan" "Afghanistan" ...
 $ Year                           : int  2015 2014 2013 2012 2011 2010 2009
2008 2007 2006 ...
 $ Status                         : chr  "Developing" "Developing"
"Developing" "Developing" ...
 $ Life.expectancy                : num  65 59.9 59.9 59.5 59.2 58.8 58.6
58.1 57.5 57.3 ...
 $ Adult.Mortality                : int  263 271 268 272 275 279 281 287 295
295 ...
 $ infant.deaths                  : int  62 64 66 69 71 74 77 80 82 84 ...
 $ Alcohol                        : num  0.01 0.01 0.01 0.01 0.01 0.01 0.01
0.03 0.02 0.03 ...
 $ percentage.expenditure         : num  71.3 73.5 73.2 78.2 7.1 ...
 $ Hepatitis.B                    : int  65 62 64 67 68 66 63 64 63 64 ...
 $ Measles                        : int  1154 492 430 2787 3013 1989 2861
1599 1141 1990 ...
 $ BMI                            : num  19.1 18.6 18.1 17.6 17.2 16.7 16.2
15.7 15.2 14.7 ...
 $ under.five.deaths              : int  83 86 89 93 97 102 106 110 113 116
...
 $ Polio                          : int  6 58 62 67 68 66 63 64 63 58 ...
 $ Total.expenditure              : num  8.16 8.18 8.13 8.52 7.87 9.2 9.42
8.33 6.73 7.43 ...
 $ Diphtheria                     : int  65 62 64 67 68 66 63 64 63 58 ...
 $ HIV.AIDS                       : num  0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
```

```
0.1 ...
 $ GDP                            : num  584.3 612.7 631.7 670 63.5 ...
 $ Population                     : num  33736494 327582 31731688 3696958
2978599 ...
 $ thinness..1.19.years           : num  17.2 17.5 17.7 17.9 18.2 18.4 18.6
18.8 19 19.2 ...
 $ thinness.5.9.years             : num  17.3 17.5 17.7 18 18.2 18.4 18.7
18.9 19.1 19.3 ...
 $ Income.composition.of.resources: num  0.479 0.476 0.47 0.463 0.454 0.448
0.434 0.433 0.415 0.405 ...
 $ Schooling                      : num  10.1 10 9.9 9.8 9.5 9.2 8.9 8.7 8.4
8.1 ...
```

Checking Observations

```
dim(WHO)
```

```
[1] 2938   22
```

Checking for Missing Values

```
# Assuming WHO is your dataframe
null_count <- colSums(is.na(WHO))
total_rows <- nrow(WHO)

# Calculate percentage of null values
percentage_null <- round((null_count / total_rows) * 100, 2)

# Create a dataframe to display results
null_summary <- data.frame(Null_Count = null_count, Percentage_Null =
percentage_null)

# Print the summary
print(null_summary)
```

```
                                 Null_Count Percentage_Null
Country                                   0            0.00
Year                                      0            0.00
Status                                    0            0.00
Life.expectancy                          10            0.34
Adult.Mortality                          10            0.34
infant.deaths                             0            0.00
Alcohol                                 194            6.60
percentage.expenditure                    0            0.00
Hepatitis.B                             553           18.82
Measles                                   0            0.00
BMI                                      34            1.16
under.five.deaths                         0            0.00
Polio                                    19            0.65
Total.expenditure                       226            7.69
Diphtheria                               19            0.65
```

```
HIV.AIDS                                0            0.00
GDP                                   448           15.25
Population                            652           22.19
thinness..1.19.years                   34            1.16
thinness.5.9.years                     34            1.16
Income.composition.of.resources       167            5.68
Schooling                             163            5.55
```

```r
for (col in names(WHO)) {
  if (is.numeric(WHO[[col]]) || is.logical(WHO[[col]])) {
    col_mean <- mean(WHO[[col]], na.rm = TRUE)
    WHO[[col]][is.na(WHO[[col]])] <- col_mean
  }
}

# Assuming WHO is your dataframe
null_count <- colSums(is.na(WHO))
total_rows <- nrow(WHO)

# Calculate percentage of null values
percentage_null <- round((null_count / total_rows) * 100, 2)

# Create a dataframe to display results
null_summary <- data.frame(Null_Count = null_count, Percentage_Null =
percentage_null)

# Print the summary
print(null_summary)
```

```
                              Null_Count Percentage_Null
Country                                0               0
Year                                   0               0
Status                                 0               0
Life.expectancy                        0               0
Adult.Mortality                        0               0
infant.deaths                          0               0
Alcohol                                0               0
percentage.expenditure                 0               0
Hepatitis.B                            0               0
Measles                                0               0
BMI                                    0               0
under.five.deaths                      0               0
Polio                                  0               0
Total.expenditure                      0               0
Diphtheria                             0               0
HIV.AIDS                               0               0
GDP                                    0               0
Population                             0               0
thinness..1.19.years                   0               0
thinness.5.9.years                     0               0
```

```
Income.composition.of.resources          0               0
Schooling                                0               0
```

Checking for Duplicate Values

```
sum(WHO[duplicated(WHO), ])
```

```
[1] 0
```

```
str(WHO)
```

```
'data.frame':   2938 obs. of  22 variables:
 $ Country                        : chr  "Afghanistan" "Afghanistan"
"Afghanistan" "Afghanistan" ...
 $ Year                           : num  2015 2014 2013 2012 2011 ...
 $ Status                         : chr  "Developing" "Developing"
"Developing" "Developing" ...
 $ Life.expectancy                : num  65 59.9 59.9 59.5 59.2 58.8 58.6
58.1 57.5 57.3 ...
 $ Adult.Mortality                : num  263 271 268 272 275 279 281 287 295
295 ...
 $ infant.deaths                  : num  62 64 66 69 71 74 77 80 82 84 ...
 $ Alcohol                        : num  0.01 0.01 0.01 0.01 0.01 0.01 0.01
0.03 0.02 0.03 ...
 $ percentage.expenditure         : num  71.3 73.5 73.2 78.2 7.1 ...
 $ Hepatitis.B                    : num  65 62 64 67 68 66 63 64 63 64 ...
 $ Measles                        : num  1154 492 430 2787 3013 ...
 $ BMI                            : num  19.1 18.6 18.1 17.6 17.2 16.7 16.2
15.7 15.2 14.7 ...
 $ under.five.deaths              : num  83 86 89 93 97 102 106 110 113 116
...
 $ Polio                          : num  6 58 62 67 68 66 63 64 63 58 ...
 $ Total.expenditure              : num  8.16 8.18 8.13 8.52 7.87 9.2 9.42
8.33 6.73 7.43 ...
 $ Diphtheria                     : num  65 62 64 67 68 66 63 64 63 58 ...
 $ HIV.AIDS                       : num  0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.1 ...
 $ GDP                            : num  584.3 612.7 631.7 670 63.5 ...
 $ Population                     : num  33736494 327582 31731688 3696958
2978599 ...
 $ thinness..1.19.years           : num  17.2 17.5 17.7 17.9 18.2 18.4 18.6
18.8 19 19.2 ...
 $ thinness.5.9.years             : num  17.3 17.5 17.7 18 18.2 18.4 18.7
18.9 19.1 19.3 ...
 $ Income.composition.of.resources: num  0.479 0.476 0.47 0.463 0.454 0.448
0.434 0.433 0.415 0.405 ...
 $ Schooling                      : num  10.1 10 9.9 9.8 9.5 9.2 8.9 8.7 8.4
8.1 ...
```

```
head(WHO)
```

```
      Country Year     Status Life.expectancy Adult.Mortality infant.deaths
1 Afghanistan 2015 Developing            65.0             263            62
2 Afghanistan 2014 Developing            59.9             271            64
3 Afghanistan 2013 Developing            59.9             268            66
4 Afghanistan 2012 Developing            59.5             272            69
5 Afghanistan 2011 Developing            59.2             275            71
6 Afghanistan 2010 Developing            58.8             279            74
  Alcohol percentage.expenditure Hepatitis.B Measles  BMI under.five.deaths
1    0.01              71.279624          65    1154 19.1                83
2    0.01              73.523582          62     492 18.6                86
3    0.01              73.219243          64     430 18.1                89
4    0.01              78.184215          67    2787 17.6                93
5    0.01               7.097109          68    3013 17.2                97
6    0.01              79.679367          66    1989 16.7               102
  Polio Total.expenditure Diphtheria HIV.AIDS       GDP Population
1     6              8.16         65      0.1 584.25921   33736494
2    58              8.18         62      0.1 612.69651     327582
3    62              8.13         64      0.1 631.74498   31731688
4    67              8.52         67      0.1 669.95900    3696958
5    68              7.87         68      0.1  63.53723    2978599
6    66              9.20         66      0.1 553.32894    2883167
  thinness..1.19.years thinness.5.9.years Income.composition.of.resources
1                 17.2                17.3                           0.479
2                 17.5                17.5                           0.476
3                 17.7                17.7                           0.470
4                 17.9                18.0                           0.463
5                 18.2                18.2                           0.454
6                 18.4                18.4                           0.448
  Schooling
1      10.1
2      10.0
3       9.9
4       9.8
5       9.5
6       9.2
```

```
WHO <- WHO[, c(2,4:22,1,3)]
```

```
str(WHO)
```

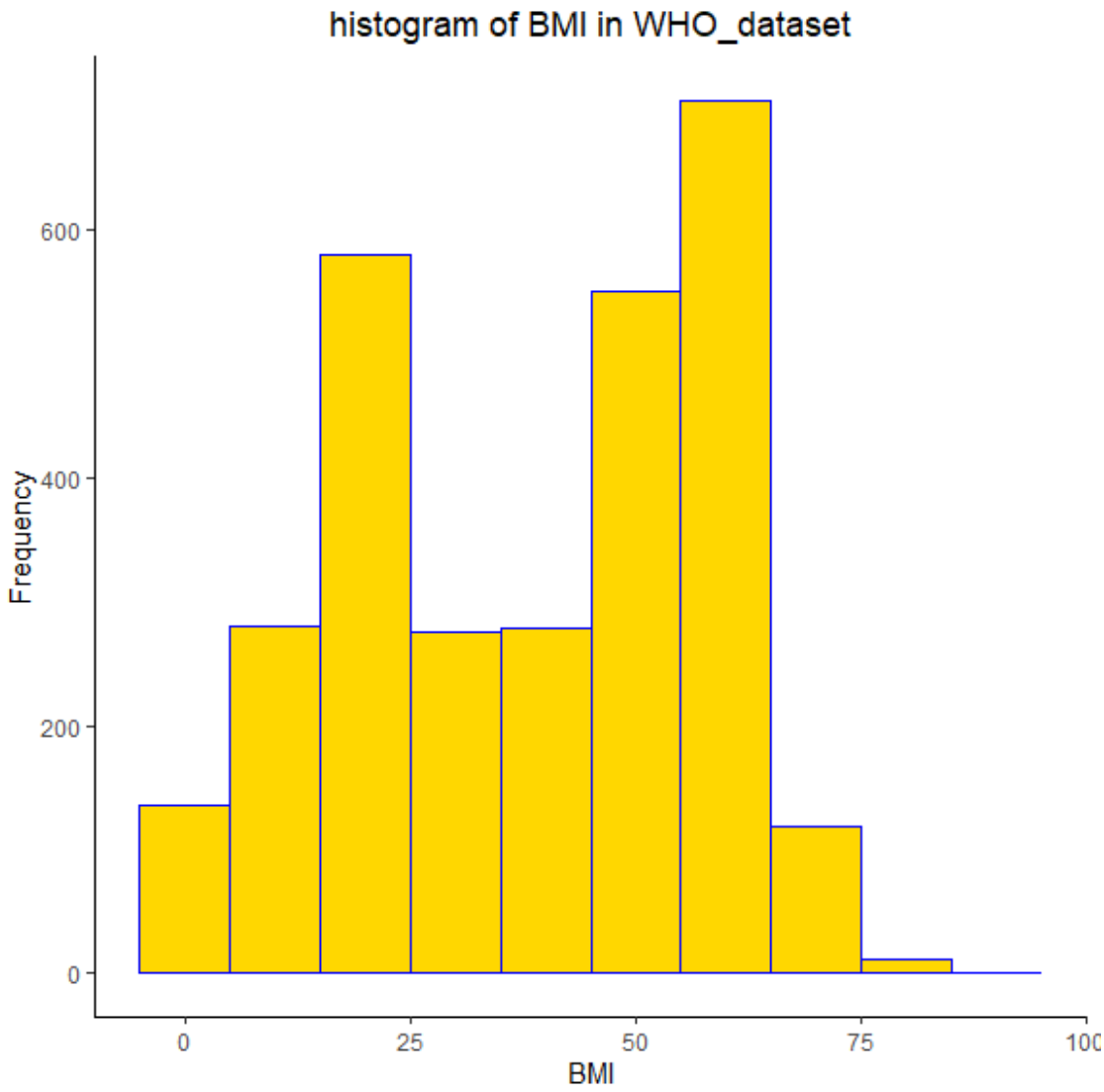
```
'data.frame':   2938 obs. of  22 variables:
 $ Year                           : num  2015 2014 2013 2012 2011 ...
 $ Life.expectancy                : num  65 59.9 59.9 59.5 59.2 58.8 58.6
58.1 57.5 57.3 ...
 $ Adult.Mortality                : num  263 271 268 272 275 279 281 287 295
295 ...
 $ infant.deaths                  : num  62 64 66 69 71 74 77 80 82 84 ...
 $ Alcohol                        : num  0.01 0.01 0.01 0.01 0.01 0.01 0.01
0.03 0.02 0.03 ...
 $ percentage.expenditure         : num  71.3 73.5 73.2 78.2 7.1 ...
```

```
 $ Hepatitis.B                    : num  65 62 64 67 68 66 63 64 63 64 ...
 $ Measles                        : num  1154 492 430 2787 3013 ...
 $ BMI                            : num  19.1 18.6 18.1 17.6 17.2 16.7 16.2
15.7 15.2 14.7 ...
 $ under.five.deaths              : num  83 86 89 93 97 102 106 110 113 116
...
 $ Polio                          : num  6 58 62 67 68 66 63 64 63 58 ...
 $ Total.expenditure              : num  8.16 8.18 8.13 8.52 7.87 9.2 9.42
8.33 6.73 7.43 ...
 $ Diphtheria                     : num  65 62 64 67 68 66 63 64 63 58 ...
 $ HIV.AIDS                       : num  0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.1 ...
 $ GDP                            : num  584.3 612.7 631.7 670 63.5 ...
 $ Population                     : num  33736494 327582 31731688 3696958
2978599 ...
 $ thinness..1.19.years           : num  17.2 17.5 17.7 17.9 18.2 18.4 18.6
18.8 19 19.2 ...
 $ thinness.5.9.years             : num  17.3 17.5 17.7 18 18.2 18.4 18.7
18.9 19.1 19.3 ...
 $ Income.composition.of.resources: num  0.479 0.476 0.47 0.463 0.454 0.448
0.434 0.433 0.415 0.405 ...
 $ Schooling                      : num  10.1 10 9.9 9.8 9.5 9.2 8.9 8.7 8.4
8.1 ...
 $ Country                        : chr  "Afghanistan" "Afghanistan"
"Afghanistan" "Afghanistan" ...
 $ Status                         : chr  "Developing" "Developing"
"Developing" "Developing" ...
```

Exploratory Data Analysis

Histogram

```
ggplot(data = WHO, aes(x = BMI )) +
  geom_histogram(binwidth = 10, fill = "gold", color = "blue") +
  labs(title = "histogram of BMI in WHO_dataset", x = "BMI", y = "Frequency")
+
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5))
```

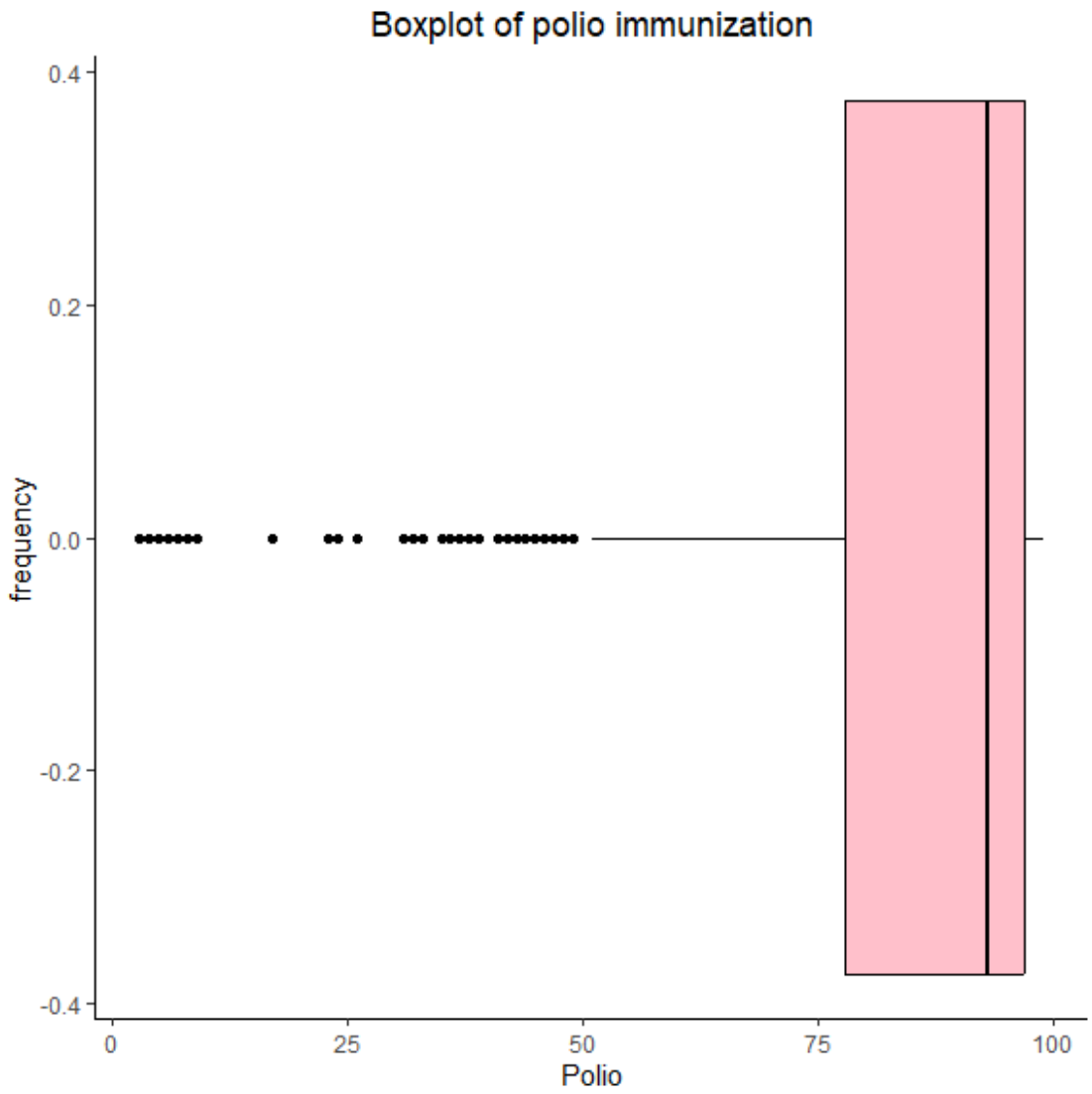



```
summary(WHO$BMI)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
   1.00   19.40   43.00   38.32   56.10   87.30 
```

Boxplot

```
ggplot(data = WHO, aes(x = Polio)) +
  geom_boxplot(fill = "pink", color = "black") +
  labs(title = "Boxplot of polio immunization", x = "Polio", y = "frequency")
+
  theme_classic() +
  theme(plot.title = element_text(hjust = 0.5))
```

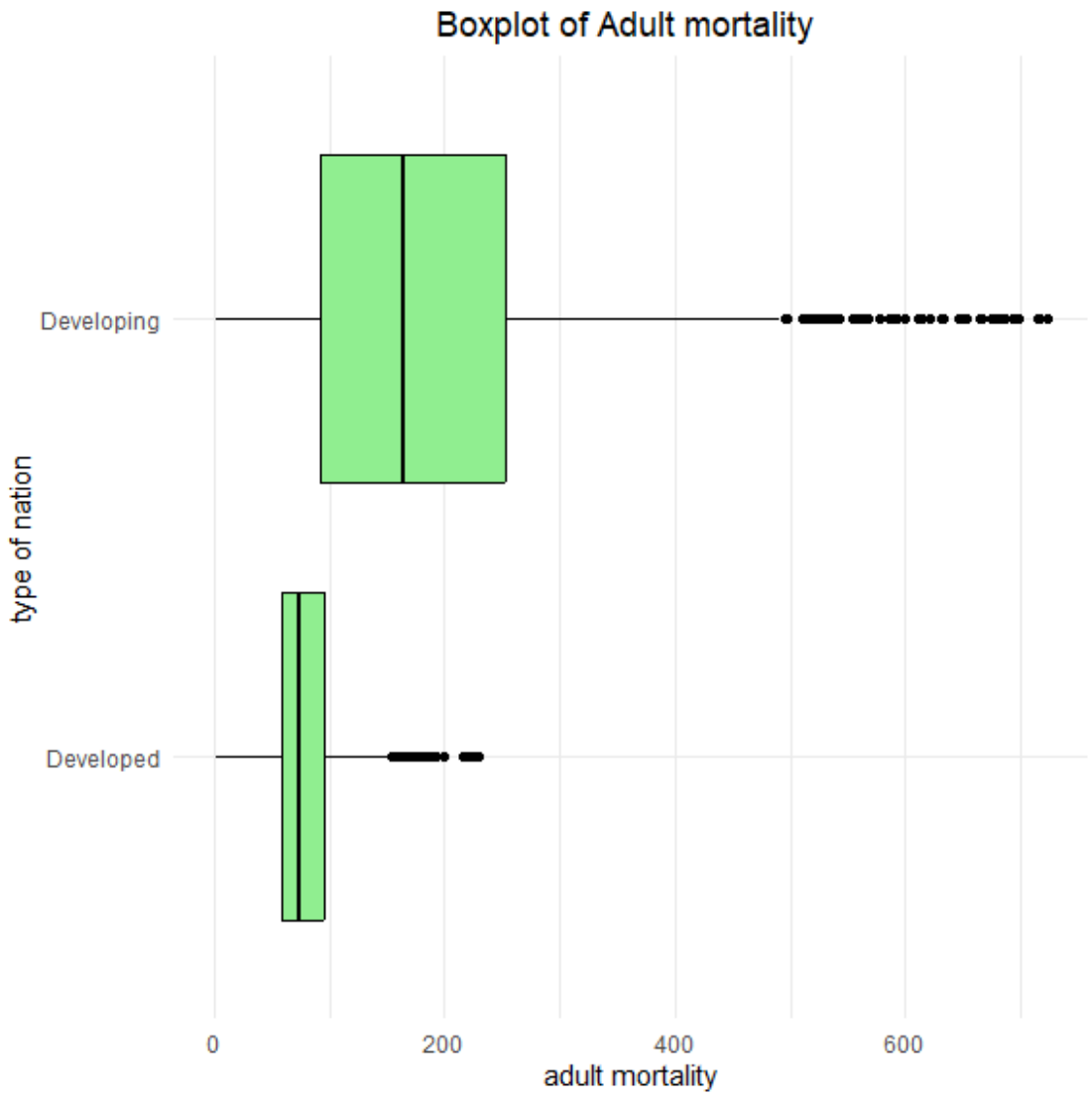



```
summary(WHO$Polio)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
   3.00   78.00   93.00   82.55   97.00   99.00 
```

Overlay Boxplot

```
ggplot(data= WHO, aes(x=Adult.Mortality, y =Status)) +
 geom_boxplot(fill = "lightgreen", color = "black") +
  labs(
   title = "Boxplot of Adult mortality ",
   x = "adult mortality",
   y = "type of nation"
 ) +
 theme_minimal() +
 theme(plot.title = element_text(hjust = 0.5))
```

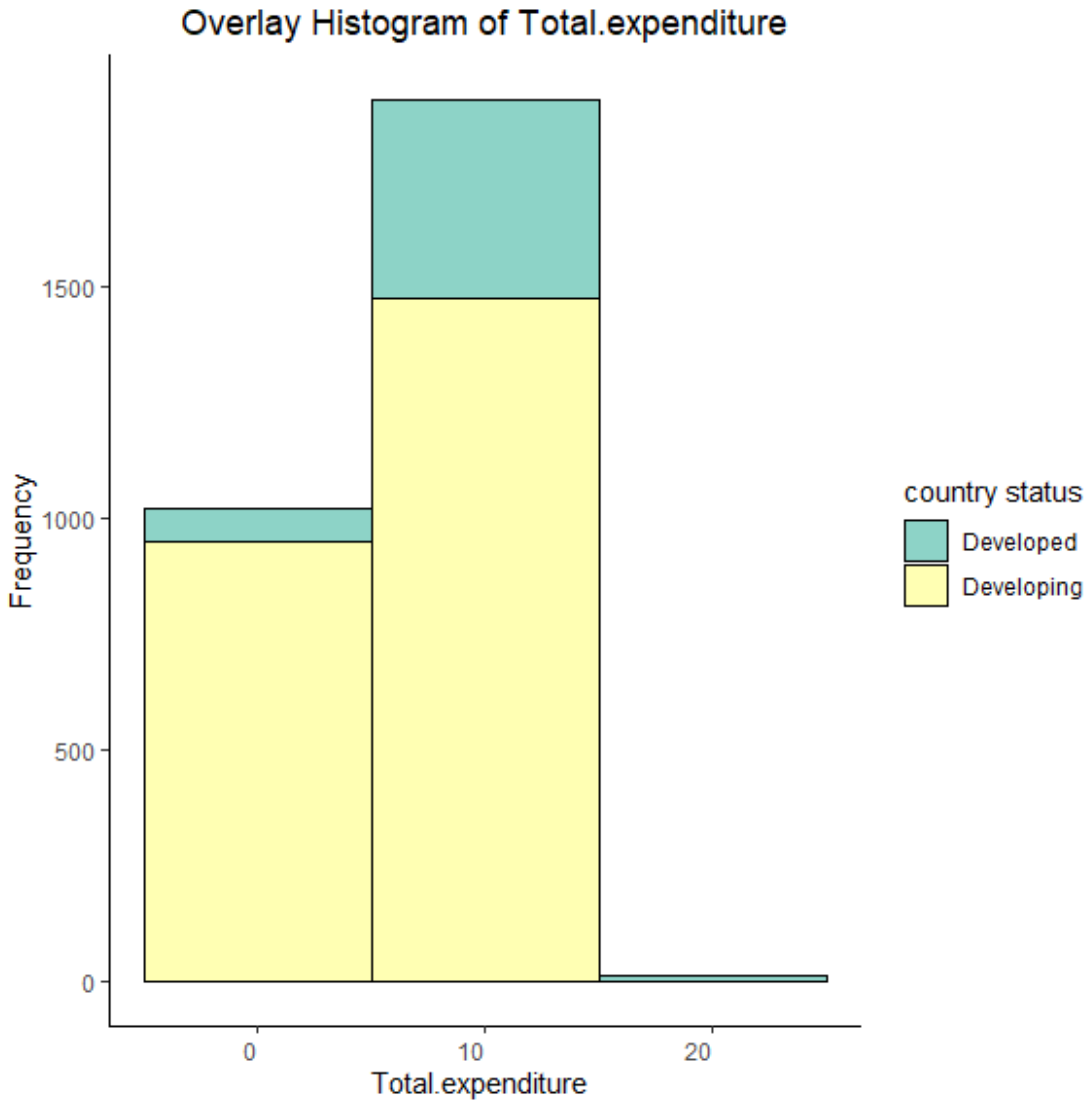


The overlay boxplot shows the adult mortality in developed and developing countries. The x axis shows the no. of adult mortality, the Y axis shows whether it is developed or developing. Developing nations are having higher adult mortality than developed nations.

Overlay Histogram

```
ggplot(data = WHO, aes(x = Total.expenditure )) +
  geom_histogram(binwidth = 10, aes(fill = Status), color = "black") +
  labs(
    x = "Total.expenditure ",
    y = "Frequency",
    fill = "country status",
    title = "Overlay Histogram of Total.expenditure"
  ) +
  theme_classic() +
  scale_fill_brewer(palette = "Set3") +
```

```
  theme(axis.text.x = element_text( hjust = 1,), plot.title =
element_text(hjust = 0.5))
```



The above overlay histogram depicts the government expenditure on health sector out of total expenditure. Clearly we can see the developed nations spending more money on health sector. Only few developed nations are spending 20% of their budget into the health sector.Most no. of nations are spending over 10% of budget and almost half of them in 0 to 10 percentage range. ## Hypotheses

```
#creating subsets
WHO_Developed <- WHO$Schooling [WHO$Status == "Developed"]

WHO_Developing <- WHO$Schooling  [WHO$Status == "Developing"]
```

Change or keep it in chunk Parameter: σ_D^2 = Variance of schooling in developed countries.

σ_{ND}^2 = Variance of schooling in developing countries.

Hypotheses:

Variance Test

```
var.test(WHO_Developed, WHO_Developing)
```

```
    F test to compare two variances

data:  WHO_Developed and WHO_Developing
F = 0.45839, num df = 511, denom df = 2425, p-value < 2.2e-16
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.4016915 0.5262423
sample estimates:
ratio of variances
         0.4583862
```

Parameter:

Change or keep it in chunk μ_D = mean of schooling in developed countries.

μ_{ND} = mean of schooling in developing.

Hypotheses:

Hypotheses Test

```
t.test(WHO_Developed, WHO_Developing, var.equal = FALSE)
```

```
    Welch Two Sample t-test

data:  WHO_Developed and WHO_Developing
t = 39.139, df = 1043.3, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 4.016370 4.440353
sample estimates:
mean of x mean of y
 15.48429  11.25592
```

Linear Regression

```
set.seed(123)

i <- sample(2, nrow(WHO), replace = TRUE, prob = c(0.8, 0.2))
WHOTraining <- WHO[i == 1, ]
```

```
WHOTest <- WHO[i == 2, ]
summary(WHOTraining)
```

```
      Year      Life.expectancy Adult.Mortality infant.deaths
 Min.   :2000   Min.   :36.30   Min.   :  1.0   Min.   :   0.00
 1st Qu.:2003   1st Qu.:63.20   1st Qu.: 74.0   1st Qu.:   0.00
 Median :2008   Median :72.00   Median :144.0   Median :   3.00
 Mean   :2008   Mean   :69.29   Mean   :164.3   Mean   :  28.99
 3rd Qu.:2012   3rd Qu.:75.80   3rd Qu.:226.5   3rd Qu.:  22.00
 Max.   :2015   Max.   :89.00   Max.   :723.0   Max.   :1700.00
    Alcohol       percentage.expenditure  Hepatitis.B       Measles
 Min.   : 0.010   Min.   :    0.00       Min.   : 1.00   Min.   :     0
 1st Qu.: 1.095   1st Qu.:    5.32       1st Qu.:80.94   1st Qu.:     0
 Median : 4.140   Median :   67.78       Median :87.00   Median :    17
 Mean   : 4.626   Mean   :  743.33       Mean   :81.11   Mean   :  2272
 3rd Qu.: 7.445   3rd Qu.:  439.53       3rd Qu.:96.00   3rd Qu.:   362
 Max.   :17.310   Max.   :19479.91       Max.   :99.00   Max.   :212183
      BMI        under.five.deaths     Polio       Total.expenditure
 Min.   : 1.00   Min.   :   0.00   Min.   : 3.00   Min.   : 0.370
 1st Qu.:19.50   1st Qu.:   0.00   1st Qu.:78.00   1st Qu.: 4.365
 Median :42.80   Median :   4.00   Median :93.00   Median : 5.938
 Mean   :38.37   Mean   :  40.34   Mean   :82.45   Mean   : 5.926
 3rd Qu.:56.10   3rd Qu.:  27.50   3rd Qu.:97.00   3rd Qu.: 7.325
 Max.   :87.30   Max.   :2300.00   Max.   :99.00   Max.   :17.600
   Diphtheria       HIV.AIDS           GDP              Population
 Min.   : 2.00   Min.   : 0.100   Min.   :     1.68   Min.   :3.600e+01
 1st Qu.:78.00   1st Qu.: 0.100   1st Qu.:   583.90   1st Qu.:4.422e+05
 Median :93.00   Median : 0.100   Median :  3129.22   Median :3.773e+06
 Mean   :82.36   Mean   : 1.707   Mean   :  7547.19   Mean   :1.294e+07
 3rd Qu.:97.00   3rd Qu.: 0.800   3rd Qu.:  7483.16   3rd Qu.:1.275e+07
 Max.   :99.00   Max.   :50.600   Max.   :119172.74   Max.   :1.294e+09
 thinness..1.19.years thinness.5.9.years Income.composition.of.resources
 Min.   : 0.100       Min.   : 0.100     Min.   :0.0000
 1st Qu.: 1.600       1st Qu.: 1.500     1st Qu.:0.5060
 Median : 3.400       Median : 3.400     Median :0.6640
 Mean   : 4.802       Mean   : 4.834     Mean   :0.6293
 3rd Qu.: 7.100       3rd Qu.: 7.100     3rd Qu.:0.7740
 Max.   :27.400       Max.   :28.400     Max.   :0.9480
   Schooling       Country             Status
 Min.   : 0.00   Length:2359        Length:2359
 1st Qu.:10.30   Class :character   Class :character
 Median :12.10   Mode  :character   Mode  :character
 Mean   :11.99
 3rd Qu.:14.20
 Max.   :20.70
```

```
summary(WHOTest)
```

```
      Year      Life.expectancy Adult.Mortality infant.deaths
 Min.   :2000   Min.   :42.30   Min.   :  1.0   Min.   :   0.00
```

```
 1st Qu.:2004   1st Qu.:62.95   1st Qu.: 75.0   1st Qu.:   0.00
 Median :2007   Median :72.00   Median :141.0   Median :   3.00
 Mean   :2008   Mean   :68.96   Mean   :166.7   Mean   :  35.67
 3rd Qu.:2012   3rd Qu.:75.30   3rd Qu.:228.0   3rd Qu.:  23.00
 Max.   :2015   Max.   :89.00   Max.   :675.0   Max.   :1800.00
    Alcohol       percentage.expenditure  Hepatitis.B       Measles
 Min.   : 0.010   Min.   :    0.000      Min.   : 2.00   Min.   :     0
 1st Qu.: 1.080   1st Qu.:    3.675      1st Qu.:80.94   1st Qu.:     0
 Median : 4.160   Median :   59.188      Median :87.00   Median :    14
 Mean   : 4.509   Mean   :  717.560      Mean   :80.27   Mean   :  3019
 3rd Qu.: 7.215   3rd Qu.:  450.049      3rd Qu.:96.00   3rd Qu.:   334
 Max.   :17.870   Max.   :18822.867      Max.   :99.00   Max.   :182485
      BMI        under.five.deaths     Polio       Total.expenditure
 Min.   : 1.90   Min.   :   0.00   Min.   : 4.00   Min.   : 1.120
 1st Qu.:19.10   1st Qu.:   0.00   1st Qu.:79.00   1st Qu.: 4.390
 Median :43.50   Median :   4.00   Median :93.00   Median : 5.900
 Mean   :38.12   Mean   :  48.94   Mean   :82.95   Mean   : 5.989
 3rd Qu.:55.80   3rd Qu.:  28.00   3rd Qu.:97.00   3rd Qu.: 7.385
 Max.   :75.70   Max.   :2500.00   Max.   :99.00   Max.   :17.200
   Diphtheria       HIV.AIDS           GDP              Population
 Min.   : 4.00   Min.   : 0.100   Min.   :    4.61   Min.   :3.400e+01
 1st Qu.:79.00   1st Qu.: 0.100   1st Qu.:  555.23   1st Qu.:3.307e+05
 Median :92.00   Median : 0.100   Median : 2932.32   Median :3.176e+06
 Mean   :82.17   Mean   : 1.886   Mean   : 7222.29   Mean   :1.198e+07
 3rd Qu.:97.00   3rd Qu.: 0.700   3rd Qu.: 7483.16   3rd Qu.:1.275e+07
 Max.   :99.00   Max.   :50.300   Max.   :87998.44   Max.   :1.144e+09
 thinness..1.19.years thinness.5.9.years Income.composition.of.resources
 Min.   : 0.100       Min.   : 0.100     Min.   :0.0000
 1st Qu.: 1.700       1st Qu.: 1.650     1st Qu.:0.5005
 Median : 3.400       Median : 3.400     Median :0.6560
 Mean   : 4.993       Mean   : 5.018     Mean   :0.6202
 3rd Qu.: 7.200       3rd Qu.: 7.250     3rd Qu.:0.7630
 Max.   :27.700       Max.   :28.600     Max.   :0.9360
   Schooling       Country             Status
 Min.   : 0.00   Length:579         Length:579
 1st Qu.:10.30   Class :character   Class :character
 Median :12.10   Mode  :character   Mode  :character
 Mean   :11.99
 3rd Qu.:13.90
 Max.   :20.40
```

```
model_1 <- lm(Adult.Mortality ~ Polio , data = WHOTraining)
summary(model_1)
```

```

Call:
lm(formula = Adult.Mortality ~ Polio, data = WHOTraining)

Residuals:
    Min      1Q  Median      3Q     Max
```

```
-262.55  -77.83  -11.67   59.96  556.00

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 285.2808     8.9030   32.04   <2e-16 ***
Polio        -1.4668     0.1038  -14.13   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 118.6 on 2357 degrees of freedom
Multiple R-squared:  0.07806,   Adjusted R-squared:  0.07767
F-statistic: 199.6 on 1 and 2357 DF,  p-value: < 2.2e-16
```

```
model_2 <- lm(Adult.Mortality ~ Polio + Alcohol + Hepatitis.B , data =
WHOTraining[, 1:20])
summary(model_2)
```

```

Call:
lm(formula = Adult.Mortality ~ Polio + Alcohol + Hepatitis.B,
    data = WHOTraining[, 1:20])

Residuals:
    Min      1Q  Median      3Q     Max
-269.63  -73.33  -15.04   58.01  562.10

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 302.1205    10.5881  28.534  < 2e-16 ***
Polio        -1.2278     0.1148 -10.696  < 2e-16 ***
Alcohol      -4.6483     0.6273  -7.409 1.76e-13 ***
Hepatitis.B  -0.1855     0.1188  -1.561    0.119
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 117.2 on 2355 degrees of freedom
Multiple R-squared:  0.1001,    Adjusted R-squared:  0.09891
F-statistic: 87.28 on 3 and 2355 DF,  p-value: < 2.2e-16
```

```
intercept_only <- lm(Adult.Mortality  ~ 1, data=WHOTraining[, 1:20])

all <- lm(Adult.Mortality ~. , data = WHOTraining[, 1:20])

model_3 <- stepAIC (intercept_only, direction='forward',scope = formula(all))

Start:  AIC=22723.63
Adult.Mortality ~ 1

                                 Df Sum of Sq      RSS   AIC
+ Life.expectancy                 1  17190188 18768561 21192
```

```
+ HIV.AIDS                            1   9707875 26250873 21983
+ Income.composition.of.resources     1   6967426 28991322 22218
+ Schooling                           1   6837001 29121747 22228
+ BMI                                 1   5029112 30929636 22370
+ thinness.5.9.years                  1   3156440 32802309 22509
+ thinness..1.19.years                1   3129674 32829075 22511
+ Diphtheria                          1   2810255 33148494 22534
+ Polio                               1   2806816 33151932 22534
+ GDP                                 1   2776367 33182382 22536
+ percentage.expenditure              1   2144987 33813762 22581
+ Alcohol                             1   1460476 34498273 22628
+ Hepatitis.B                         1    749415 35209333 22676
+ Total.expenditure                   1    412948 35545800 22698
+ under.five.deaths                   1    296077 35662672 22706
+ infant.deaths                       1    198692 35760056 22713
+ Year                                1    190121 35768628 22713
+ Measles                             1     40443 35918305 22723
<none>                                             35958749 22724
+ Population                          1     10463 35948286 22725

Step:  AIC=21191.83
Adult.Mortality ~ Life.expectancy

                                   Df Sum of Sq      RSS   AIC
+ HIV.AIDS                          1    942385 17826176 21072
+ Measles                           1    252038 18516523 21162
+ Schooling                         1    244731 18523830 21163
+ Alcohol                           1    222331 18546229 21166
+ under.five.deaths                 1    205338 18563223 21168
+ infant.deaths                     1    186204 18582357 21170
+ Diphtheria                        1    135355 18633205 21177
+ Income.composition.of.resources   1    117784 18650776 21179
+ Polio                             1    100992 18667569 21181
+ Year                              1     88871 18679690 21183
+ thinness..1.19.years              1     60206 18708355 21186
+ Total.expenditure                 1     56364 18712197 21187
+ thinness.5.9.years                1     36228 18732333 21189
+ Population                        1     34705 18733855 21190
<none>                                         18768561 21192
+ GDP                               1     13506 18755055 21192
+ percentage.expenditure            1     12732 18755829 21192
+ Hepatitis.B                       1      2002 18766559 21194
+ BMI                               1      1504 18767057 21194

Step:  AIC=21072.31
Adult.Mortality ~ Life.expectancy + HIV.AIDS

                                   Df Sum of Sq      RSS   AIC
+ Measles                          1    194937 17631239 21048
```

```
+ Year                             1    126948 17699228 21057
+ under.five.deaths                1    125296 17700880 21058
+ infant.deaths                    1    111904 17714272 21060
+ Alcohol                          1     71939 17754238 21065
+ Diphtheria                       1     60202 17765974 21066
+ Schooling                        1     45524 17780652 21068
+ Polio                            1     40503 17785674 21069
+ thinness..1.19.years             1     26266 17799910 21071
+ Population                       1     20802 17805375 21072
+ Income.composition.of.resources  1     17337 17808839 21072
<none>                                       17826176 21072
+ thinness.5.9.years               1     12768 17813409 21073
+ Total.expenditure                1     11075 17815101 21073
+ BMI                              1      1894 17824282 21074
+ percentage.expenditure           1       761 17825415 21074
+ Hepatitis.B                      1       476 17825700 21074
+ GDP                              1       227 17825949 21074

Step:  AIC=21048.37
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles

                                  Df Sum of Sq      RSS   AIC
+ Year                             1    111629 17519610 21035
+ Alcohol                          1     82551 17548689 21039
+ Diphtheria                       1     48316 17582923 21044
+ Schooling                        1     47944 17583295 21044
+ Polio                            1     32233 17599006 21046
+ under.five.deaths                1     22265 17608974 21047
+ infant.deaths                    1     17048 17614191 21048
+ Income.composition.of.resources  1     16274 17614965 21048
<none>                                       17631239 21048
+ thinness..1.19.years             1      8567 17622672 21049
+ BMI                              1      7402 17623837 21049
+ Total.expenditure                1      6129 17625110 21050
+ thinness.5.9.years               1      1828 17629411 21050
+ Population                       1       887 17630352 21050
+ percentage.expenditure           1       336 17630903 21050
+ GDP                              1        78 17631161 21050
+ Hepatitis.B                      1        31 17631209 21050

Step:  AIC=21035.39
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year

                                  Df Sum of Sq      RSS   AIC
+ Alcohol                          1    107327 17412283 21023
+ Diphtheria                       1     39461 17480149 21032
+ Polio                            1     30477 17489134 21033
+ Schooling                        1     30119 17489492 21033
+ under.five.deaths                1     27296 17492315 21034
```

```
+ infant.deaths                     1     21691 17497919 21035
<none>                                       17519610 21035
+ thinness..1.19.years              1     11675 17507935 21036
+ BMI                               1      7750 17511860 21036
+ Income.composition.of.resources   1      4190 17515420 21037
+ Total.expenditure                 1      4013 17515598 21037
+ thinness.5.9.years                1      3477 17516133 21037
+ Population                        1      1845 17517765 21037
+ Hepatitis.B                       1       547 17519063 21037
+ GDP                               1       359 17519251 21037
+ percentage.expenditure            1        80 17519531 21037

Step:  AIC=21022.89
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol

                                  Df Sum of Sq      RSS   AIC
+ Diphtheria                       1     37823 17374461 21020
+ Polio                            1     29200 17383083 21021
+ under.five.deaths                1     26364 17385919 21021
+ infant.deaths                    1     19071 17393212 21022
+ BMI                              1     16637 17395647 21023
<none>                                       17412283 21023
+ Schooling                        1      5746 17406537 21024
+ percentage.expenditure           1      5156 17407127 21024
+ GDP                              1      4760 17407523 21024
+ Population                       1      1354 17410929 21025
+ thinness.5.9.years               1       923 17411360 21025
+ Hepatitis.B                      1       809 17411474 21025
+ thinness..1.19.years             1       301 17411982 21025
+ Total.expenditure                1       197 17412086 21025
+ Income.composition.of.resources  1         4 17412280 21025

Step:  AIC=21019.76
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria

                                  Df Sum of Sq      RSS   AIC
+ under.five.deaths                1   21981.6 17352479 21019
+ Hepatitis.B                      1   19531.7 17354929 21019
+ BMI                              1   17153.4 17357307 21019
+ infant.deaths                    1   15961.4 17358499 21020
<none>                                       17374461 21020
+ Polio                            1    5503.6 17368957 21021
+ Schooling                        1    4746.7 17369714 21021
+ GDP                              1    3469.5 17370991 21021
+ percentage.expenditure           1    3420.8 17371040 21021
+ Population                       1    1372.6 17373088 21022
+ thinness.5.9.years               1     779.5 17373681 21022
```

```
+ Total.expenditure                1     547.7 17373913 21022
+ thinness..1.19.years             1     354.2 17374106 21022
+ Income.composition.of.resources  1      77.9 17374383 21022

Step:  AIC=21018.77
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria + under.five.deaths

                                  Df Sum of Sq      RSS   AIC
+ infant.deaths                    1     46396 17306083 21015
+ Hepatitis.B                      1     26255 17326224 21017
+ BMI                              1     21381 17331097 21018
<none>                                         17352479 21019
+ thinness.5.9.years               1      9643 17342835 21020
+ Polio                            1      5074 17347405 21020
+ Schooling                        1      3751 17348728 21020
+ GDP                              1      3748 17348731 21020
+ percentage.expenditure           1      3424 17349055 21020
+ Population                       1      2563 17349916 21020
+ thinness..1.19.years             1      2023 17350456 21021
+ Total.expenditure                1       919 17351560 21021
+ Income.composition.of.resources  1        61 17352418 21021

Step:  AIC=21014.46
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria + under.five.deaths + infant.deaths

                                  Df Sum of Sq      RSS   AIC
+ Hepatitis.B                      1   21109.0 17284974 21014
+ BMI                              1   18916.5 17287166 21014
<none>                                         17306083 21015
+ thinness.5.9.years               1    5008.2 17301075 21016
+ Polio                            1    4336.2 17301747 21016
+ Schooling                        1    3932.9 17302150 21016
+ GDP                              1    2013.5 17304069 21016
+ percentage.expenditure           1    1976.2 17304107 21016
+ Total.expenditure                1     720.8 17305362 21016
+ thinness..1.19.years             1     552.3 17305531 21016
+ Income.composition.of.resources  1     273.3 17305810 21016
+ Population                       1      21.3 17306062 21017

Step:  AIC=21013.58
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria + under.five.deaths + infant.deaths +
    Hepatitis.B

                                  Df Sum of Sq      RSS   AIC
+ BMI                              1   19425.4 17265548 21013
<none>                                         17284974 21014
```

```
+ Polio                              1    6979.1 17277995 21015
+ thinness.5.9.years                1    6278.1 17278696 21015
+ Schooling                         1    4435.0 17280539 21015
+ percentage.expenditure            1    2721.3 17282253 21015
+ GDP                               1    2172.3 17282802 21015
+ thinness..1.19.years              1    1040.4 17283933 21015
+ Total.expenditure                 1     802.9 17284171 21016
+ Income.composition.of.resources   1     313.1 17284661 21016
+ Population                        1       6.0 17284968 21016

Step:  AIC=21012.93
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria + under.five.deaths + infant.deaths +
    Hepatitis.B + BMI

                                  Df Sum of Sq      RSS   AIC
<none>                                         17265548 21013
+ Schooling                        1    7432.0 17258116 21014
+ Polio                            1    7372.9 17258176 21014
+ percentage.expenditure           1    2971.2 17262577 21015
+ GDP                              1    1939.3 17263609 21015
+ thinness.5.9.years               1    1202.9 17264346 21015
+ Total.expenditure                1     168.2 17265380 21015
+ thinness..1.19.years             1     147.9 17265401 21015
+ Population                        1      56.0 17265492 21015
+ Income.composition.of.resources   1       7.2 17265541 21015

model_3$anova

Stepwise Model Path
Analysis of Deviance Table

Initial Model:
Adult.Mortality ~ 1

Final Model:
Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles + Year +
    Alcohol + Diphtheria + under.five.deaths + infant.deaths +
    Hepatitis.B + BMI

                   Step Df    Deviance Resid. Df Resid. Dev      AIC
1                                          2358   35958749 22723.63
2    + Life.expectancy  1 17190187.76      2357   18768561 21191.83
3           + HIV.AIDS  1   942384.56      2356   17826176 21072.31
4            + Measles  1   194937.12      2355   17631239 21048.37
5               + Year  1   111628.77      2354   17519610 21035.39
6            + Alcohol  1   107327.27      2353   17412283 21022.89
7         + Diphtheria  1    37822.58      2352   17374461 21019.76
8  + under.five.deaths  1    21981.64      2351   17352479 21018.77
9      + infant.deaths  1    46396.00      2350   17306083 21014.46
```

```
10       + Hepatitis.B  1    21109.03      2349   17284974 21013.58
11               + BMI  1    19425.43      2348   17265548 21012.93
```

```r
summary(model_3)
```

```

Call:
lm(formula = Adult.Mortality ~ Life.expectancy + HIV.AIDS + Measles +
    Year + Alcohol + Diphtheria + under.five.deaths + infant.deaths +
    Hepatitis.B + BMI, data = WHOTraining[, 1:20])

Residuals:
    Min      1Q  Median      3Q     Max
-578.01   -8.21   16.31   41.30  237.56

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2.679e+03  7.872e+02  -3.403 0.000677 ***
Life.expectancy   -8.648e+00  3.141e-01 -27.533  < 2e-16 ***
HIV.AIDS           4.209e+00  4.452e-01   9.456  < 2e-16 ***
Measles           -6.131e-04  1.957e-04  -3.133 0.001753 **
Year               1.708e+00  3.932e-01   4.342 1.47e-05 ***
Alcohol            2.245e+00  5.170e-01   4.342 1.47e-05 ***
Diphtheria         2.415e-01  1.002e-01   2.411 0.015979 *
under.five.deaths -3.692e-01  1.477e-01  -2.500 0.012489 *
infant.deaths      4.578e-01  1.992e-01   2.299 0.021616 *
Hepatitis.B       -1.616e-01  9.425e-02  -1.715 0.086548 .
BMI               -1.762e-01  1.084e-01  -1.625 0.104224
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 85.75 on 2348 degrees of freedom
Multiple R-squared:  0.5199,    Adjusted R-squared:  0.5178
F-statistic: 254.2 on 10 and 2348 DF,  p-value: < 2.2e-16
```

```r
model_4<- stepAIC (all, direction='backward')
```

```
Start:  AIC=21026.67
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Total.expenditure + Diphtheria + HIV.AIDS + GDP +
    Population + thinness..1.19.years + thinness.5.9.years +
    Income.composition.of.resources + Schooling

                                Df Sum of Sq      RSS   AIC
- GDP                            1        23 17234428 21025
- Total.expenditure              1        36 17234441 21025
- Population                     1        80 17234485 21025
- percentage.expenditure         1       971 17235376 21025
```

```
- Income.composition.of.resources  1      3311 17237716 21025
- Polio                            1      7522 17241928 21026
- thinness..1.19.years             1      8455 17242860 21026
- thinness.5.9.years               1      9095 17243500 21026
- Schooling                        1     11246 17245651 21026
<none>                                       17234405 21027
- BMI                              1     17969 17252374 21027
- Diphtheria                       1     19453 17253858 21027
- Hepatitis.B                      1     26395 17260801 21028
- infant.deaths                    1     31415 17265820 21029
- under.five.deaths                1     39143 17273548 21030
- Measles                          1     72908 17307313 21035
- Alcohol                          1    104271 17338676 21039
- Year                             1    124631 17359036 21042
- HIV.AIDS                         1    577160 17811565 21102
- Life.expectancy                  1   3752173 20986578 21489

Step:  AIC=21024.67
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Total.expenditure + Diphtheria + HIV.AIDS + Population +
    thinness..1.19.years + thinness.5.9.years +
Income.composition.of.resources +
    Schooling

                                  Df Sum of Sq      RSS   AIC
- Total.expenditure                1        44 17234472 21023
- Population                       1        82 17234509 21023
- percentage.expenditure           1      2920 17237348 21023
- Income.composition.of.resources  1      3288 17237716 21023
- Polio                            1      7574 17242002 21024
- thinness..1.19.years             1      8438 17242865 21024
- thinness.5.9.years               1      9077 17243505 21024
- Schooling                        1     11235 17245663 21024
<none>                                       17234428 21025
- BMI                              1     17960 17252388 21025
- Diphtheria                       1     19430 17253858 21025
- Hepatitis.B                      1     26403 17260831 21026
- infant.deaths                    1     31398 17265826 21027
- under.five.deaths                1     39128 17273555 21028
- Measles                          1     72890 17307318 21033
- Alcohol                          1    104271 17338699 21037
- Year                             1    125944 17360372 21040
- HIV.AIDS                         1    577469 17811897 21100
- Life.expectancy                  1   3763894 20998322 21489

Step:  AIC=21022.68
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
```

```
+
    Polio + Diphtheria + HIV.AIDS + Population + thinness..1.19.years +
    thinness.5.9.years + Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- Population                       1        84 17234556 21021
- percentage.expenditure           1      2987 17237458 21021
- Income.composition.of.resources  1      3245 17237717 21021
- Polio                            1      7573 17242044 21022
- thinness..1.19.years             1      8440 17242912 21022
- thinness.5.9.years               1      9195 17243666 21022
- Schooling                        1     11214 17245685 21022
<none>                                       17234472 21023
- BMI                              1     18193 17252664 21023
- Diphtheria                       1     19386 17253858 21023
- Hepatitis.B                      1     26403 17260875 21024
- infant.deaths                    1     31365 17265836 21025
- under.five.deaths                1     39096 17273568 21026
- Measles                          1     72868 17307340 21031
- Alcohol                          1    106957 17341429 21035
- Year                             1    126605 17361076 21038
- HIV.AIDS                         1    580725 17815196 21099
- Life.expectancy                  1   3770693 21005165 21487

Step:  AIC=21020.69
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years
+
    Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- percentage.expenditure           1      2994 17237550 21019
- Income.composition.of.resources  1      3252 17237808 21019
- Polio                            1      7559 17242115 21020
- thinness..1.19.years             1      8392 17242948 21020
- thinness.5.9.years               1      9142 17243698 21020
- Schooling                        1     11263 17245819 21020
<none>                                       17234556 21021
- BMI                              1     18139 17252695 21021
- Diphtheria                       1     19471 17254027 21021
- Hepatitis.B                      1     26451 17261007 21022
- infant.deaths                    1     33609 17268165 21023
- under.five.deaths                1     40613 17275169 21024
- Measles                          1     73065 17307621 21029
- Alcohol                          1    107067 17341623 21033
- Year                             1    126675 17361231 21036
- HIV.AIDS                         1    580754 17815310 21097
- Life.expectancy                  1   3770646 21005202 21485
```

```
Step:  AIC=21019.1
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years +
    Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- Income.composition.of.resources  1      3873 17241423 21018
- Polio                            1      7719 17245269 21018
- thinness..1.19.years             1      8555 17246105 21018
- thinness.5.9.years               1      9461 17247012 21018
- Schooling                        1     10842 17248393 21019
<none>                                       17237550 21019
- BMI                              1     17548 17255098 21020
- Diphtheria                       1     19779 17257330 21020
- Hepatitis.B                      1     25634 17263184 21021
- infant.deaths                    1     35199 17272749 21022
- under.five.deaths                1     42391 17279941 21023
- Measles                          1     73027 17310577 21027
- Alcohol                          1    104256 17341807 21031
- Year                             1    128050 17365600 21035
- HIV.AIDS                         1    577810 17815360 21095
- Life.expectancy                  1   3909547 21147097 21499

Step:  AIC=21017.63
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years +
    Schooling

                      Df Sum of Sq      RSS   AIC
- Schooling            1      7004 17248427 21017
- Polio                1      7577 17249000 21017
- thinness..1.19.years 1      8258 17249681 21017
- thinness.5.9.years   1      9405 17250828 21017
<none>                             17241423 21018
- BMI                  1     17944 17259367 21018
- Diphtheria           1     19491 17260914 21018
- Hepatitis.B          1     25146 17266569 21019
- infant.deaths        1     34015 17275438 21020
- under.five.deaths    1     41230 17282654 21021
- Measles              1     72122 17313545 21026
- Alcohol              1    103351 17344774 21030
- Year                 1    124356 17365780 21033
- HIV.AIDS             1    574405 17815828 21093
- Life.expectancy      1   4133987 21375411 21523

Step:  AIC=21016.59
```

```
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years

                       Df Sum of Sq      RSS   AIC
- Polio                 1      7814 17256242 21016
- thinness..1.19.years  1      8614 17257041 21016
- thinness.5.9.years    1      9542 17257969 21016
<none>                              17248427 21017
- BMI                   1     15609 17264037 21017
- Diphtheria            1     19600 17268027 21017
- Hepatitis.B           1     24422 17272849 21018
- infant.deaths         1     34102 17282529 21019
- under.five.deaths     1     41400 17289827 21020
- Measles               1     70511 17318939 21024
- Alcohol               1    128299 17376726 21032
- Year                  1    139840 17388267 21034
- HIV.AIDS              1    644487 17892914 21101
- Life.expectancy       1   5360309 22608736 21653

Step:  AIC=21015.65
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS + thinness..1.19.years + thinness.5.9.years

                       Df Sum of Sq      RSS   AIC
- thinness..1.19.years  1      8104 17264346 21015
- thinness.5.9.years    1      9159 17265401 21015
<none>                              17256242 21016
- BMI                   1     15161 17271403 21016
- Hepatitis.B           1     21601 17277843 21017
- infant.deaths         1     35284 17291526 21019
- Diphtheria            1     42390 17298632 21019
- under.five.deaths     1     42745 17298987 21020
- Measles               1     70897 17327139 21023
- Alcohol               1    129222 17385464 21031
- Year                  1    138081 17394323 21033
- HIV.AIDS              1    654614 17910856 21102
- Life.expectancy       1   5456009 22712251 21662

Step:  AIC=21014.76
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS + thinness.5.9.years

                      Df Sum of Sq      RSS   AIC
- thinness.5.9.years  1      1203 17265548 21013
- BMI                 1     14350 17278696 21015
<none>                            17264346 21015
```

```
- Hepatitis.B         1     22127 17286473 21016
- infant.deaths       1     36500 17300845 21018
- Diphtheria          1     42689 17307034 21019
- under.five.deaths   1     44418 17308763 21019
- Measles             1     70585 17334931 21022
- Alcohol             1    136621 17400966 21031
- Year                1    138581 17402926 21032
- HIV.AIDS            1    654965 17919311 21101
- Life.expectancy     1   5448272 22712617 21660

Step:  AIC=21012.93
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS

                     Df Sum of Sq      RSS   AIC
<none>                             17265548 21013
- BMI                 1     19425 17284974 21014
- Hepatitis.B         1     21618 17287166 21014
- infant.deaths       1     38851 17304400 21016
- Diphtheria          1     42749 17308297 21017
- under.five.deaths   1     45956 17311504 21017
- Measles             1     72167 17337715 21021
- Alcohol             1    138636 17404184 21030
- Year                1    138658 17404206 21030
- HIV.AIDS            1    657452 17923001 21099
- Life.expectancy     1   5574181 22839730 21671
```

```
summary(model_4)
```

```

Call:
lm(formula = Adult.Mortality ~ Year + Life.expectancy + infant.deaths +
    Alcohol + Hepatitis.B + Measles + BMI + under.five.deaths +
    Diphtheria + HIV.AIDS, data = WHOTraining[, 1:20])

Residuals:
    Min      1Q  Median      3Q     Max
-578.01   -8.21   16.31   41.30  237.56

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2.679e+03  7.872e+02  -3.403 0.000677 ***
Year               1.708e+00  3.932e-01   4.342 1.47e-05 ***
Life.expectancy   -8.648e+00  3.141e-01 -27.533  < 2e-16 ***
infant.deaths      4.578e-01  1.992e-01   2.299 0.021616 *
Alcohol            2.245e+00  5.170e-01   4.342 1.47e-05 ***
Hepatitis.B       -1.616e-01  9.425e-02  -1.715 0.086548 .
Measles           -6.131e-04  1.957e-04  -3.133 0.001753 **
BMI               -1.762e-01  1.084e-01  -1.625 0.104224
```

```
under.five.deaths -3.692e-01  1.477e-01  -2.500 0.012489 *
Diphtheria         2.415e-01  1.002e-01   2.411 0.015979 *
HIV.AIDS           4.209e+00  4.452e-01   9.456  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 85.75 on 2348 degrees of freedom
Multiple R-squared:  0.5199,    Adjusted R-squared:  0.5178
F-statistic: 254.2 on 10 and 2348 DF,  p-value: < 2.2e-16
```

```
model_5<- stepAIC (all, direction='both')
```

```
Start:  AIC=21026.67
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Total.expenditure + Diphtheria + HIV.AIDS + GDP +
    Population + thinness..1.19.years + thinness.5.9.years +
    Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- GDP                              1        23 17234428 21025
- Total.expenditure                1        36 17234441 21025
- Population                       1        80 17234485 21025
- percentage.expenditure           1       971 17235376 21025
- Income.composition.of.resources  1      3311 17237716 21025
- Polio                            1      7522 17241928 21026
- thinness..1.19.years             1      8455 17242860 21026
- thinness.5.9.years               1      9095 17243500 21026
- Schooling                        1     11246 17245651 21026
<none>                                       17234405 21027
- BMI                              1     17969 17252374 21027
- Diphtheria                       1     19453 17253858 21027
- Hepatitis.B                      1     26395 17260801 21028
- infant.deaths                    1     31415 17265820 21029
- under.five.deaths                1     39143 17273548 21030
- Measles                          1     72908 17307313 21035
- Alcohol                          1    104271 17338676 21039
- Year                             1    124631 17359036 21042
- HIV.AIDS                         1    577160 17811565 21102
- Life.expectancy                  1   3752173 20986578 21489

Step:  AIC=21024.67
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Total.expenditure + Diphtheria + HIV.AIDS + Population +
    thinness..1.19.years + thinness.5.9.years +
Income.composition.of.resources +
    Schooling
```

```
                                  Df Sum of Sq      RSS   AIC
- Total.expenditure                1        44 17234472 21023
- Population                       1        82 17234509 21023
- percentage.expenditure           1      2920 17237348 21023
- Income.composition.of.resources  1      3288 17237716 21023
- Polio                            1      7574 17242002 21024
- thinness..1.19.years             1      8438 17242865 21024
- thinness.5.9.years               1      9077 17243505 21024
- Schooling                        1     11235 17245663 21024
<none>                                       17234428 21025
- BMI                              1     17960 17252388 21025
- Diphtheria                       1     19430 17253858 21025
- Hepatitis.B                      1     26403 17260831 21026
+ GDP                              1        23 17234405 21027
- infant.deaths                    1     31398 17265826 21027
- under.five.deaths                1     39128 17273555 21028
- Measles                          1     72890 17307318 21033
- Alcohol                          1    104271 17338699 21037
- Year                             1    125944 17360372 21040
- HIV.AIDS                         1    577469 17811897 21100
- Life.expectancy                  1   3763894 20998322 21489

Step:  AIC=21022.68
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Diphtheria + HIV.AIDS + Population + thinness..1.19.years +
    thinness.5.9.years + Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- Population                       1        84 17234556 21021
- percentage.expenditure           1      2987 17237458 21021
- Income.composition.of.resources  1      3245 17237717 21021
- Polio                            1      7573 17242044 21022
- thinness..1.19.years             1      8440 17242912 21022
- thinness.5.9.years               1      9195 17243666 21022
- Schooling                        1     11214 17245685 21022
<none>                                       17234472 21023
- BMI                              1     18193 17252664 21023
- Diphtheria                       1     19386 17253858 21023
- Hepatitis.B                      1     26403 17260875 21024
+ Total.expenditure                1        44 17234428 21025
+ GDP                              1        30 17234441 21025
- infant.deaths                    1     31365 17265836 21025
- under.five.deaths                1     39096 17273568 21026
- Measles                          1     72868 17307340 21031
- Alcohol                          1    106957 17341429 21035
- Year                             1    126605 17361076 21038
- HIV.AIDS                         1    580725 17815196 21099
```

```
- Life.expectancy                   1   3770693 21005165 21487

Step:  AIC=21020.69
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    percentage.expenditure + Hepatitis.B + Measles + BMI + under.five.deaths
+
    Polio + Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years
+
    Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- percentage.expenditure           1      2994 17237550 21019
- Income.composition.of.resources  1      3252 17237808 21019
- Polio                            1      7559 17242115 21020
- thinness..1.19.years             1      8392 17242948 21020
- thinness.5.9.years               1      9142 17243698 21020
- Schooling                        1     11263 17245819 21020
<none>                                       17234556 21021
- BMI                              1     18139 17252695 21021
- Diphtheria                       1     19471 17254027 21021
- Hepatitis.B                      1     26451 17261007 21022
+ Population                       1        84 17234472 21023
+ Total.expenditure                1        47 17234509 21023
+ GDP                              1        32 17234524 21023
- infant.deaths                    1     33609 17268165 21023
- under.five.deaths                1     40613 17275169 21024
- Measles                          1     73065 17307621 21029
- Alcohol                          1    107067 17341623 21033
- Year                             1    126675 17361231 21036
- HIV.AIDS                         1    580754 17815310 21097
- Life.expectancy                  1   3770646 21005202 21485

Step:  AIC=21019.1
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years +
    Income.composition.of.resources + Schooling

                                  Df Sum of Sq      RSS   AIC
- Income.composition.of.resources  1      3873 17241423 21018
- Polio                            1      7719 17245269 21018
- thinness..1.19.years             1      8555 17246105 21018
- thinness.5.9.years               1      9461 17247012 21018
- Schooling                        1     10842 17248393 21019
<none>                                       17237550 21019
- BMI                              1     17548 17255098 21020
- Diphtheria                       1     19779 17257330 21020
- Hepatitis.B                      1     25634 17263184 21021
+ percentage.expenditure           1      2994 17234556 21021
```

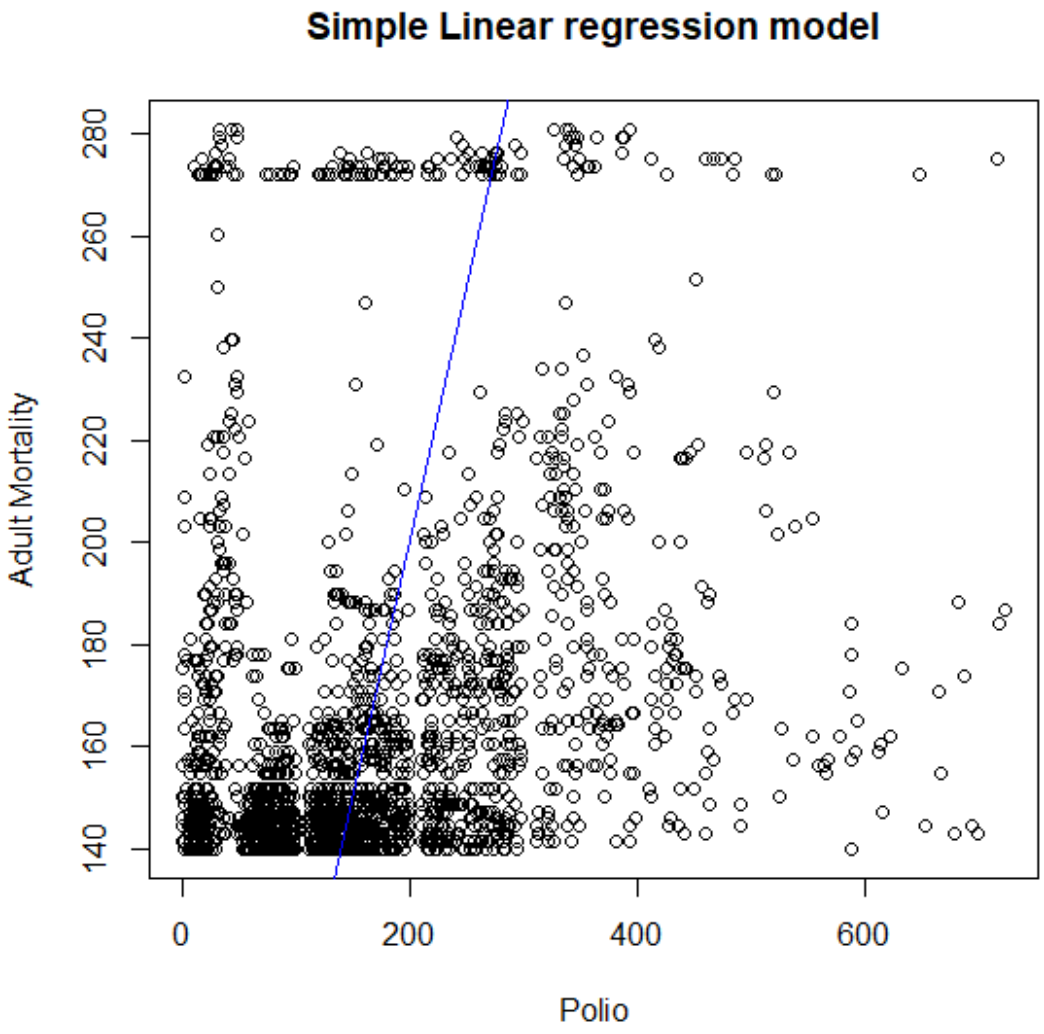
```
+ GDP                                1      1969 17235581 21021
+ Total.expenditure                  1       116 17237435 21021
+ Population                         1        92 17237458 21021
- infant.deaths                      1     35199 17272749 21022
- under.five.deaths                  1     42391 17279941 21023
- Measles                            1     73027 17310577 21027
- Alcohol                            1    104256 17341807 21031
- Year                               1    128050 17365600 21035
- HIV.AIDS                           1    577810 17815360 21095
- Life.expectancy                    1   3909547 21147097 21499

Step:  AIC=21017.63
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years +
    Schooling

                                   Df Sum of Sq      RSS   AIC
- Schooling                         1      7004 17248427 21017
- Polio                             1      7577 17249000 21017
- thinness..1.19.years              1      8258 17249681 21017
- thinness.5.9.years                1      9405 17250828 21017
<none>                                         17241423 21018
- BMI                               1     17944 17259367 21018
- Diphtheria                        1     19491 17260914 21018
- Hepatitis.B                       1     25146 17266569 21019
+ Income.composition.of.resources   1      3873 17237550 21019
+ percentage.expenditure            1      3615 17237808 21019
+ GDP                               1      2721 17238702 21019
+ Population                        1       101 17241322 21020
+ Total.expenditure                 1        21 17241402 21020
- infant.deaths                     1     34015 17275438 21020
- under.five.deaths                 1     41230 17282654 21021
- Measles                           1     72122 17313545 21026
- Alcohol                           1    103351 17344774 21030
- Year                              1    124356 17365780 21033
- HIV.AIDS                          1    574405 17815828 21093
- Life.expectancy                   1   4133987 21375411 21523

Step:  AIC=21016.59
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Polio +
    Diphtheria + HIV.AIDS + thinness..1.19.years + thinness.5.9.years

                                   Df Sum of Sq      RSS   AIC
- Polio                             1      7814 17256242 21016
- thinness..1.19.years              1      8614 17257041 21016
- thinness.5.9.years                1      9542 17257969 21016
<none>                                         17248427 21017
```

```
- BMI                                1     15609 17264037 21017
- Diphtheria                        1     19600 17268027 21017
+ Schooling                         1      7004 17241423 21018
- Hepatitis.B                       1     24422 17272849 21018
+ percentage.expenditure            1      2608 17245820 21018
+ GDP                               1      1770 17246658 21018
+ Population                        1       141 17248287 21019
+ Total.expenditure                 1        64 17248364 21019
+ Income.composition.of.resources   1        35 17248393 21019
- infant.deaths                     1     34102 17282529 21019
- under.five.deaths                 1     41400 17289827 21020
- Measles                           1     70511 17318939 21024
- Alcohol                           1    128299 17376726 21032
- Year                              1    139840 17388267 21034
- HIV.AIDS                          1    644487 17892914 21101
- Life.expectancy                   1   5360309 22608736 21653

Step:  AIC=21015.65
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS + thinness..1.19.years + thinness.5.9.years

                                  Df Sum of Sq      RSS   AIC
- thinness..1.19.years             1      8104 17264346 21015
- thinness.5.9.years               1      9159 17265401 21015
<none>                                        17256242 21016
- BMI                              1     15161 17271403 21016
+ Polio                            1      7814 17248427 21017
- Hepatitis.B                      1     21601 17277843 21017
+ Schooling                        1      7242 17249000 21017
+ percentage.expenditure           1      2731 17253511 21017
+ GDP                              1      1686 17254556 21017
+ Population                       1       122 17256120 21018
+ Total.expenditure                1        66 17256175 21018
+ Income.composition.of.resources  1        17 17256225 21018
- infant.deaths                    1     35284 17291526 21019
- Diphtheria                       1     42390 17298632 21019
- under.five.deaths                1     42745 17298987 21020
- Measles                          1     70897 17327139 21023
- Alcohol                          1    129222 17385464 21031
- Year                             1    138081 17394323 21033
- HIV.AIDS                         1    654614 17910856 21102
- Life.expectancy                  1   5456009 22712251 21662

Step:  AIC=21014.76
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS + thinness.5.9.years
```

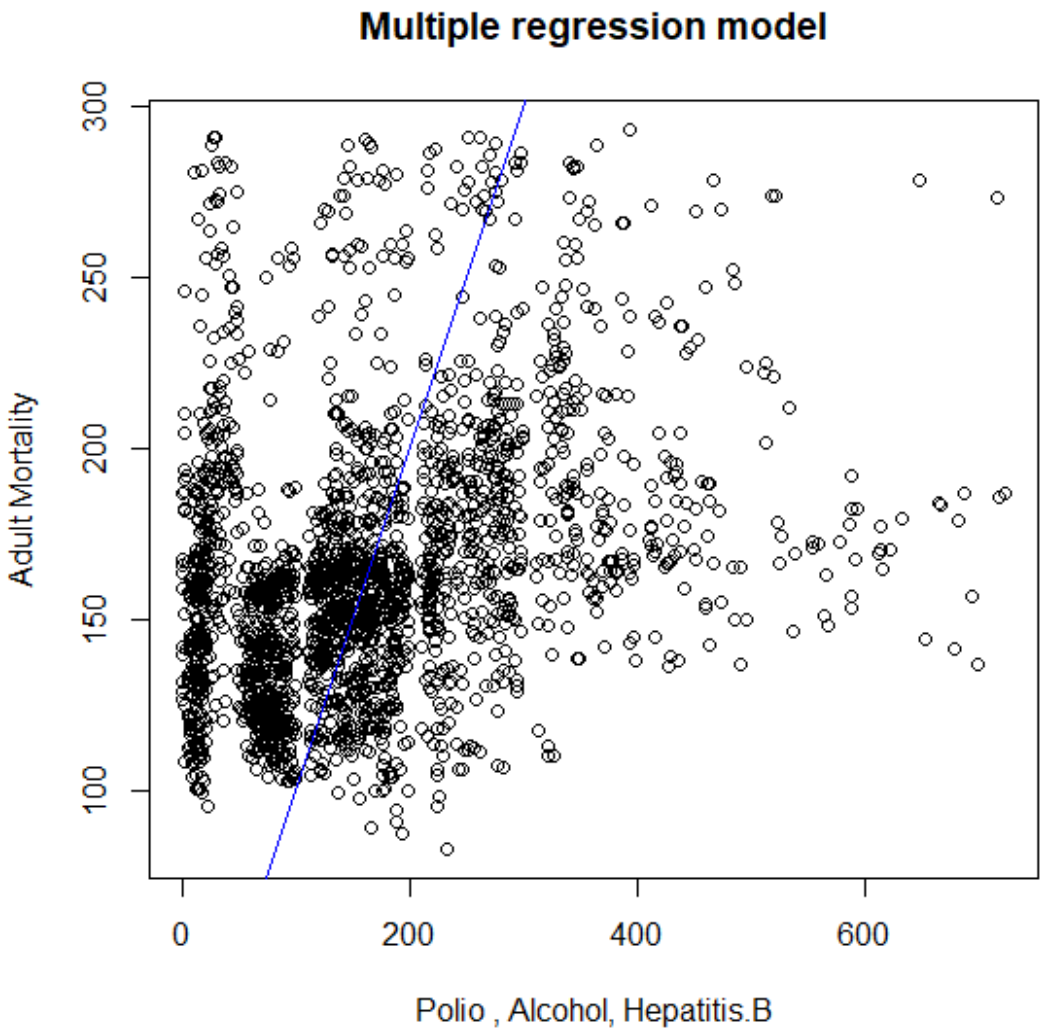
```
                                  Df Sum of Sq      RSS   AIC
- thinness.5.9.years               1      1203 17265548 21013
- BMI                              1     14350 17278696 21015
<none>                                       17264346 21015
+ thinness..1.19.years             1      8104 17256242 21016
+ Schooling                        1      7585 17256761 21016
+ Polio                            1      7304 17257041 21016
- Hepatitis.B                      1     22127 17286473 21016
+ percentage.expenditure           1      2831 17261514 21016
+ GDP                              1      1847 17262499 21017
+ Total.expenditure                1        75 17264270 21017
+ Population                       1        64 17264281 21017
+ Income.composition.of.resources  1         1 17264344 21017
- infant.deaths                    1     36500 17300845 21018
- Diphtheria                       1     42689 17307034 21019
- under.five.deaths                1     44418 17308763 21019
- Measles                          1     70585 17334931 21022
- Alcohol                          1    136621 17400966 21031
- Year                             1    138581 17402926 21032
- HIV.AIDS                         1    654965 17919311 21101
- Life.expectancy                  1   5448272 22712617 21660

Step:  AIC=21012.93
Adult.Mortality ~ Year + Life.expectancy + infant.deaths + Alcohol +
    Hepatitis.B + Measles + BMI + under.five.deaths + Diphtheria +
    HIV.AIDS

                                  Df Sum of Sq      RSS   AIC
<none>                                       17265548 21013
- BMI                              1     19425 17284974 21014
- Hepatitis.B                      1     21618 17287166 21014
+ Schooling                        1      7432 17258116 21014
+ Polio                            1      7373 17258176 21014
+ percentage.expenditure           1      2971 17262577 21015
+ GDP                              1      1939 17263609 21015
+ thinness.5.9.years               1      1203 17264346 21015
+ Total.expenditure                1       168 17265380 21015
+ thinness..1.19.years             1       148 17265401 21015
+ Population                       1        56 17265492 21015
+ Income.composition.of.resources  1         7 17265541 21015
- infant.deaths                    1     38851 17304400 21016
- Diphtheria                       1     42749 17308297 21017
- under.five.deaths                1     45956 17311504 21017
- Measles                          1     72167 17337715 21021
- Alcohol                          1    138636 17404184 21030
- Year                             1    138658 17404206 21030
- HIV.AIDS                         1    657452 17923001 21099
- Life.expectancy                  1   5574181 22839730 21671

summary(model_5)
```

```
Call:
lm(formula = Adult.Mortality ~ Year + Life.expectancy + infant.deaths +
    Alcohol + Hepatitis.B + Measles + BMI + under.five.deaths +
    Diphtheria + HIV.AIDS, data = WHOTraining[, 1:20])

Residuals:
    Min      1Q  Median      3Q     Max
-578.01   -8.21   16.31   41.30  237.56

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2.679e+03  7.872e+02  -3.403 0.000677 ***
Year               1.708e+00  3.932e-01   4.342 1.47e-05 ***
Life.expectancy   -8.648e+00  3.141e-01 -27.533  < 2e-16 ***
infant.deaths      4.578e-01  1.992e-01   2.299 0.021616 *
Alcohol            2.245e+00  5.170e-01   4.342 1.47e-05 ***
Hepatitis.B       -1.616e-01  9.425e-02  -1.715 0.086548 .
Measles           -6.131e-04  1.957e-04  -3.133 0.001753 **
BMI               -1.762e-01  1.084e-01  -1.625 0.104224
under.five.deaths -3.692e-01  1.477e-01  -2.500 0.012489 *
Diphtheria         2.415e-01  1.002e-01   2.411 0.015979 *
HIV.AIDS           4.209e+00  4.452e-01   9.456  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 85.75 on 2348 degrees of freedom
Multiple R-squared:  0.5199,    Adjusted R-squared:  0.5178
F-statistic: 254.2 on 10 and 2348 DF,  p-value: < 2.2e-16
```

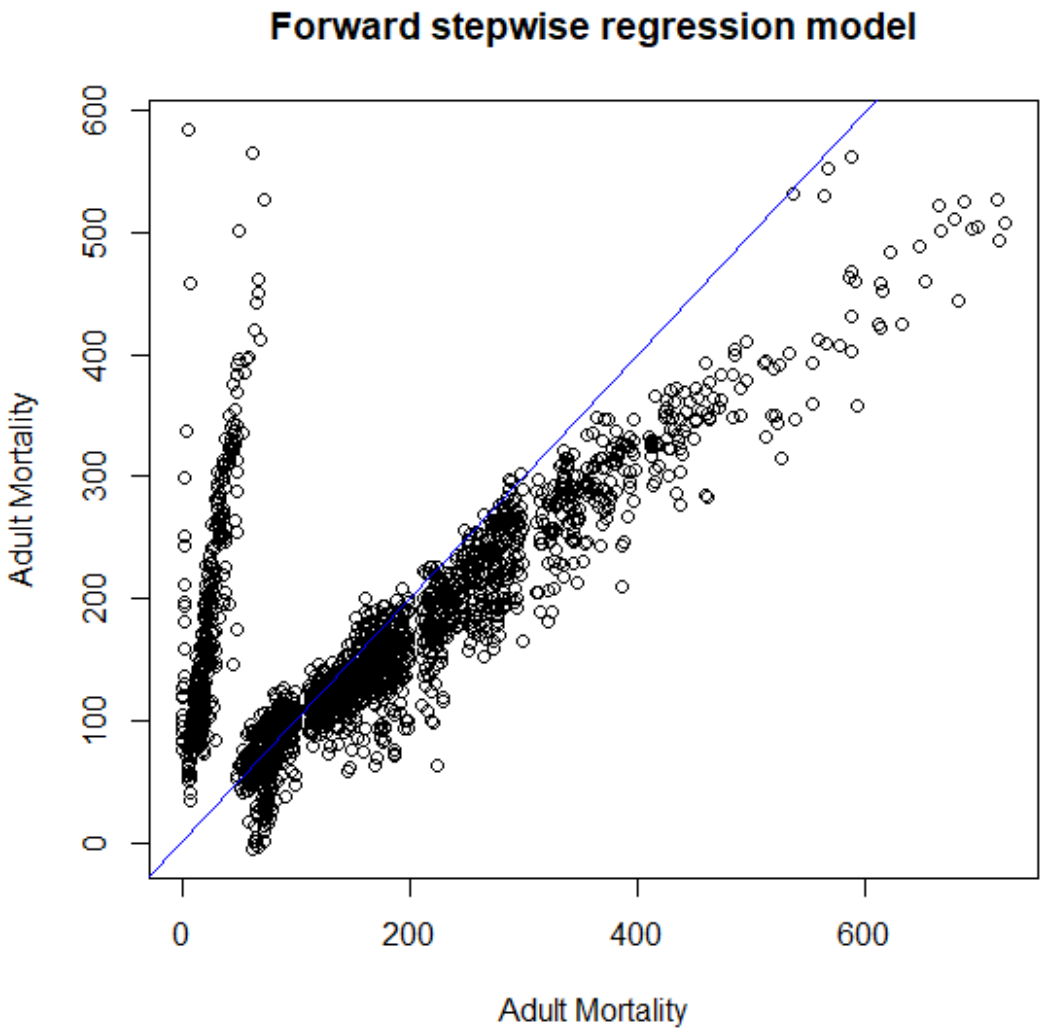
```r
plot(WHOTraining$Adult.Mortality, fitted(model_1),xlab = "Polio", ylab =
"Adult Mortality", main = "Simple Linear regression model")

abline(0, 1, col = "blue")
```



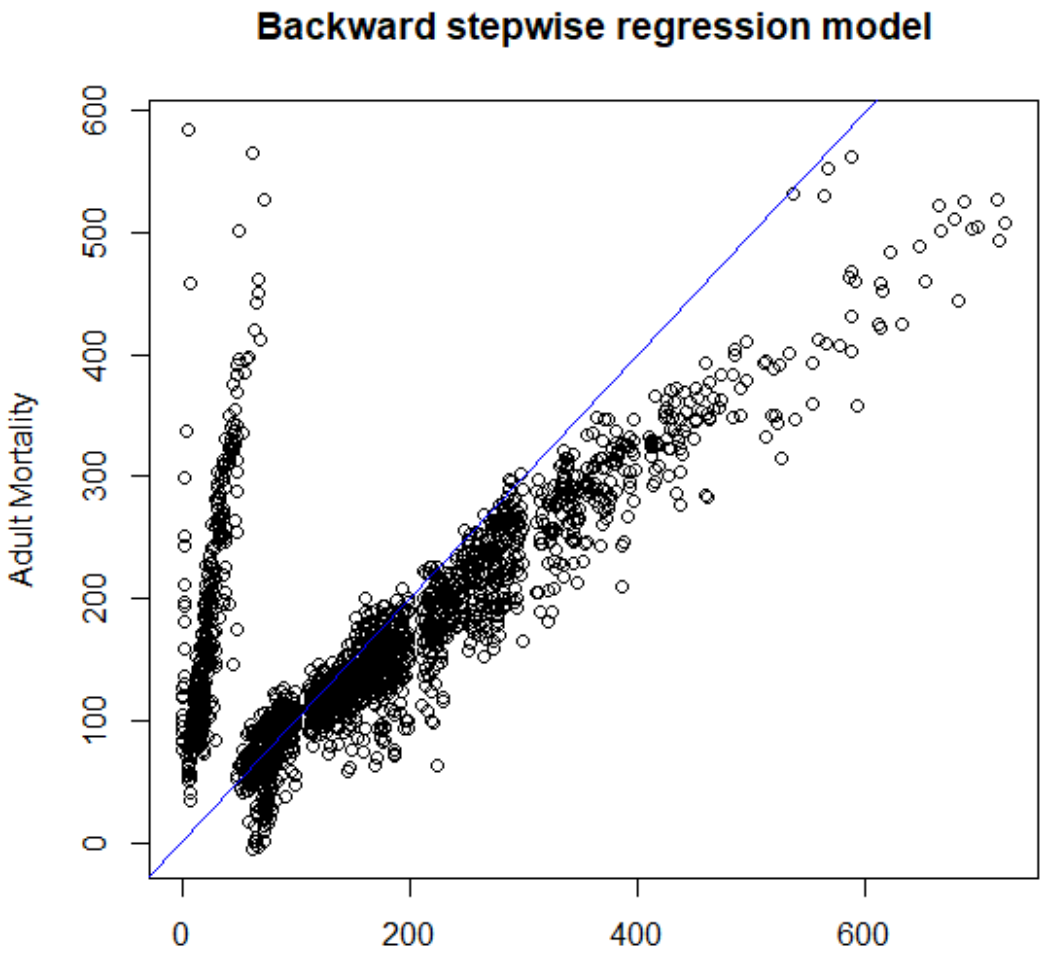

```
plot(WHOTraining$Adult.Mortality,fitted(model_2),xlab = "Polio , Alcohol,
Hepatitis.B", ylab = "Adult Mortality", main = "Multiple regression model")

abline(0, 1, col = "blue")
```



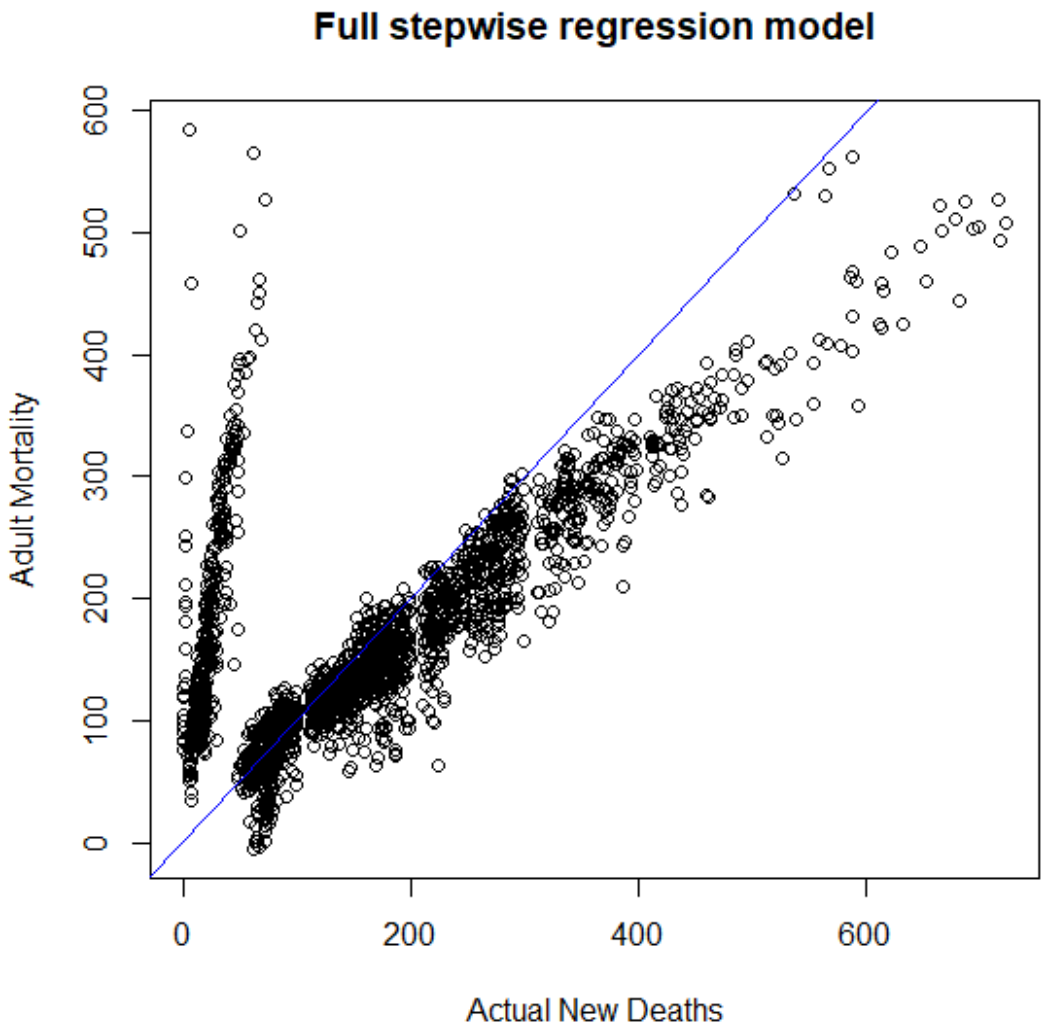

```
plot(WHOTraining$Adult.Mortality,fitted(model_3),xlab = "Adult Mortality",
ylab = "Adult Mortality", main = "Forward stepwise regression model")

abline(0, 1, col = "blue")
```




```
plot(WHOTraining$Adult.Mortality,fitted(model_4),xlab = "", ylab = "Adult
Mortality", main = "Backward stepwise regression model")

abline(0, 1, col = "blue")
```




```
plot(WHOTraining$Adult.Mortality,fitted(model_5),xlab = "Actual New Deaths",
ylab = "Adult Mortality", main = "Full stepwise regression model")

abline(0, 1, col = "blue")
```




```r
aic_values <- c(
  AIC(model_1),
  AIC(model_2),
  AIC(model_3),
  AIC(model_4),
  AIC(model_5)
)

#Add

adjusted_r_squared <- c(
  summary(model_1)$adj.r.squared,
  summary(model_2)$adj.r.squared,
  summary(model_3)$adj.r.squared,
   summary(model_4)$adj.r.squared,
   summary(model_5)$adj.r.squared
```

```
)

#change
predict_model_1 <- predict(model_1, newdata = WHOTest)
predict_model_2 <- predict(model_2, newdata = WHOTest)
predict_model_3 <- predict(model_3, newdata = WHOTest)
predict_model_4 <- predict(model_4, newdata = WHOTest)
predict_model_5 <- predict(model_5, newdata = WHOTest)

mae_values <- c(
  MAE(y_pred = predict_model_1, y_true = WHOTest$Adult.Mortality),
  MAE(y_pred = predict_model_2, y_true = WHOTest$Adult.Mortality),
  MAE(y_pred = predict_model_3, y_true = WHOTest$Adult.Mortality),
   MAE(y_pred = predict_model_4, y_true = WHOTest$Adult.Mortality),
    MAE(y_pred = predict_model_5, y_true = WHOTest$Adult.Mortality)
)

mse_values <- c(
  MSE(y_pred = predict_model_1, y_true = WHOTest$Adult.Mortality),
  MSE(y_pred = predict_model_2, y_true = WHOTest$Adult.Mortality),
  MSE(y_pred = predict_model_3, y_true = WHOTest$Adult.Mortality),
    MSE(y_pred = predict_model_4, y_true = WHOTest$Adult.Mortality),
    MSE(y_pred = predict_model_5, y_true = WHOTest$Adult.Mortality)
)

#change
estimation_table <- data.frame(
  Model = c("Simple Linear Regression","multiple", "Forward Stepwise",
"Backward Stepwise","full"),
  AIC = aic_values,
  Adjusted_R_Squared = adjusted_r_squared,
  MAE = mae_values,
  MSE = mse_values
)
estimation_table
```

```
                    Model      AIC Adjusted_R_Squared      MAE       MSE
1 Simple Linear Regression 29230.46         0.07766541 91.66538 15034.531
2                 multiple 29177.48         0.09891147 90.90056 14898.332
3         Forward Stepwise 27709.48         0.51780632 55.99464  7234.001
4        Backward Stepwise 27709.48         0.51780632 55.99464  7234.001
5                     full 27709.48         0.51780632 55.99464  7234.001
```