This is an R HTML document. When you click the **Knit HTML** button a web page will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#For plotting the scatter density plots
library(GGally)
```

```
## Loading required package: ggplot2
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
##
## Attaching package: 'GGally'
```

```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

```
##############reading the dataset################################

who <- read.csv("C:/Users/prera/Downloads/Life_Expectancy_Data.csv")
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
```

```
############Dimesion of the dataset############################

dim(who)
```

```
## [1] 2938   22
```

```
##############     TOP 10 DEVELPED & DEVELOPING Coutires      #####################

status.of.countries <- who[(who$Status %in% c("Developing") & who$Life.expectancy<55) | (who$Status %in% c("Developed") & who$Life.expectancy>85) ,]
dim(status.of.countries)
```

```
## [1] 347  22
```

```r
#View(status.of.countries)

class(status.of.countries)
```

```
## [1] "data.frame"
```

```r
head(status.of.countries)
```

```
##        Country Year     Status Life.expectancy Adult.Mortality infant.deaths
## 16 Afghanistan 2000 Developing            54.8             321            88
## 49      Angola 2015 Developing            52.4             335            66
## 50      Angola 2014 Developing            51.7             348            67
## 51      Angola 2013 Developing            51.1             355            69
## 53      Angola 2011 Developing            51.0             361            75
## 54      Angola 2010 Developing            49.6             365            78
##    Alcohol percentage.expenditure Hepatitis.B Measles  BMI under.five.deaths
## 16    0.01               10.42496          62    6532 12.2               122
## 49      NA                0.00000          64     118 23.3                98
## 50    8.33               23.96561          64   11699 22.7               101
## 51    8.10               35.95857          77    8523 22.1               105
## 53    8.06              239.89139          72    1449 21.0               115
## 54    7.80              191.65374          77    1190  2.4               121
##    Polio Total.expenditure Diphtheria HIV.AIDS       GDP Population
## 16    24              8.20         24      0.1  114.5600     293756
## 49     7                NA         64      1.9 3695.7937    2785935
## 50    68              3.31         64      2.0  479.3122    2692466
## 51    67              4.26         77      2.3  484.6169    2599834
## 53    73              3.38         71      2.5 4299.1289   24218565
## 54    81              3.39         77      2.5 3529.5348   23369131
##    thinness..1.19.years thinness.5.9.years Income.composition.of.resources
## 16                  2.3                2.5                           0.338
## 49                  8.3                8.2                           0.531
## 50                  8.5                8.3                           0.527
## 51                  8.6                8.5                           0.523
## 53                  8.9                8.8                           0.495
## 54                  9.1                9.0                           0.488
##    Schooling
## 16       5.5
## 49      11.4
## 50      11.4
## 51      11.4
## 53       9.4
## 54       9.0
```

```r
#View(status.of.countries)
WHONew<-status.of.countries
#resting the index values
row.names(WHONew) <- NULL
#View(WHONew)
dim(WHONew)
```

```
## [1] 347  22
```

```r
###################### CLEANING THE DATA ######################################
# For 347 rows running the for loop for chechking any NA values and replacing it with the mean of the
# particular country.
for(i in 1:347)
{
  if(is.na(WHONew$Alcohol[i]))
  {
    WHONew$Alcohol[i] <- with(WHONew, mean(WHONew$Alcohol[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
for(i in 1:347)
{
  if(is.na(WHONew$Hepatitis.B[i]))
  {
    WHONew$Hepatitis.B[i] <- with(WHONew, mean(WHONew$Hepatitis.B[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
for(i in 1:347)
{
  if(is.na(WHONew$Total.expenditure[i]))
  {
    WHONew$Total.expenditure[i] <- with(WHONew, mean(WHONew$Total.expenditure[Country == WHONew$Country[i]], na.rm = TRUE))
  }
}
dim(WHONew)
```

```
## [1] 347  22
```

```r
#View(WHONew)

# Deleting the Empty rows where there is no data present.
new.life<- na.omit(WHONew)
dim(new.life)
```

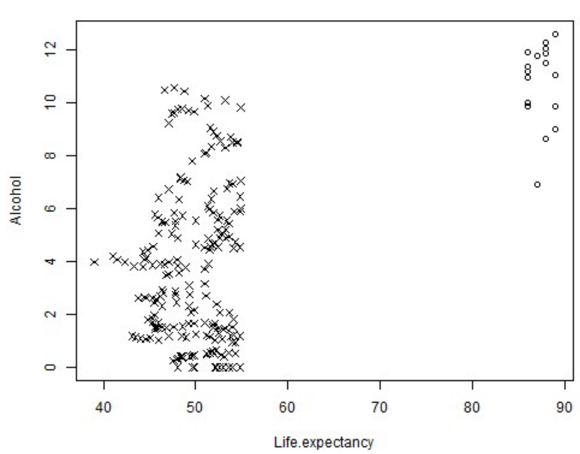
```
## [1] 223  22
```

```
#View(new.life)

############ Finding the summary or the distribution of the Dataset ##################

g<-lm(Life.expectancy~Adult.Mortality + 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, data=new.life)
summary(g)
```

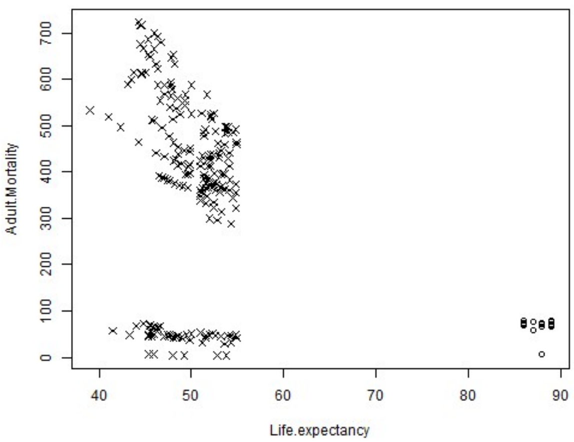
```
##
## Call:
## lm(formula = Life.expectancy ~ Adult.Mortality + infant.deaths +
##     Alcohol + percentage.expenditure + Hepatitis.B + Measles +
##     BMI + under.five.deaths + Polio + Total.expenditure + Diphtheria +
##     HIV.AIDS + GDP + Population + thinness..1.19.years + thinness.5.9.years +
##     Income.composition.of.resources + Schooling, data = new.life)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.9519 -2.8746 -0.1113  2.7742 14.2823
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     4.015e+01  1.691e+00  23.743  < 2e-16 ***
## Adult.Mortality                -7.540e-03  1.589e-03  -4.745 3.92e-06 ***
## infant.deaths                  -2.131e-01  8.926e-02  -2.387 0.017885 *
## Alcohol                         2.426e-01  1.302e-01   1.863 0.063882 .
## percentage.expenditure          2.871e-03  1.359e-03   2.112 0.035895 *
## Hepatitis.B                    -2.038e-03  1.529e-02  -0.133 0.894141
## Measles                        -3.869e-06  1.784e-05  -0.217 0.828508
## BMI                             1.235e-01  3.472e-02   3.556 0.000467 ***
## under.five.deaths               1.295e-01  5.677e-02   2.281 0.023586 *
## Polio                           1.471e-02  1.565e-02   0.940 0.348177
## Total.expenditure              -2.483e-01  1.231e-01  -2.017 0.044967 *
## Diphtheria                      3.611e-02  1.706e-02   2.116 0.035521 *
## HIV.AIDS                       -2.923e-01  2.832e-02 -10.324  < 2e-16 ***
## GDP                            -3.021e-04  1.999e-04  -1.511 0.132276
## Population                      1.259e-08  1.474e-08   0.854 0.394036
## thinness..1.19.years           -8.334e-02  1.580e-01  -0.527 0.598561
## thinness.5.9.years             -2.873e-01  1.537e-01  -1.870 0.062974 .
## Income.composition.of.resources  1.539e+01  4.660e+00   3.302 0.001134 **
## Schooling                       9.968e-01  2.459e-01   4.054 7.16e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.296 on 204 degrees of freedom
## Multiple R-squared:  0.8561, Adjusted R-squared:  0.8434
## F-statistic: 67.42 on 18 and 204 DF,  p-value: < 2.2e-16
```

You can also embed plots, for example:

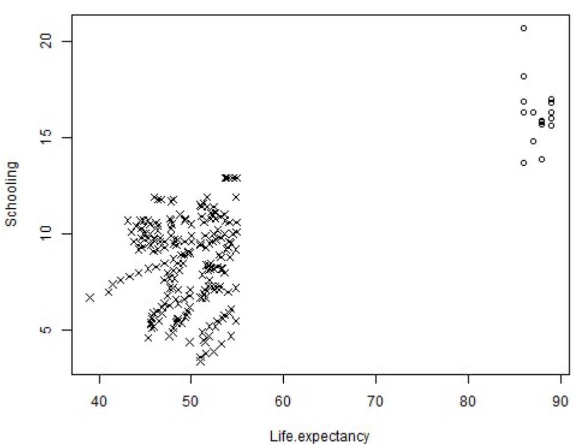
```
##################PLOTTING FOR CORREALTION#####################################
attach(new.life)

plot(Life.expectancy, Alcohol, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="Alcohol")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



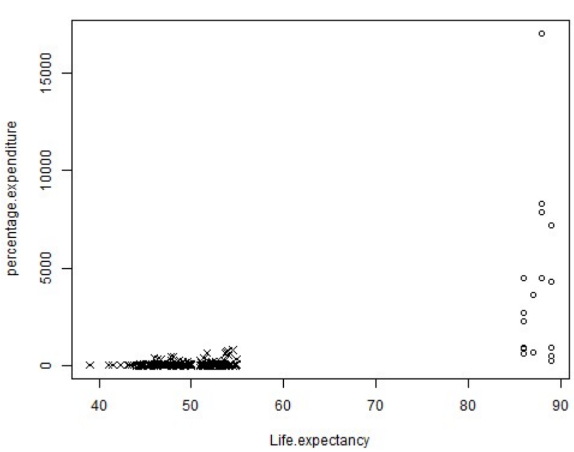

```
plot(Life.expectancy,Adult.Mortality, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="Adult.Mortality")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



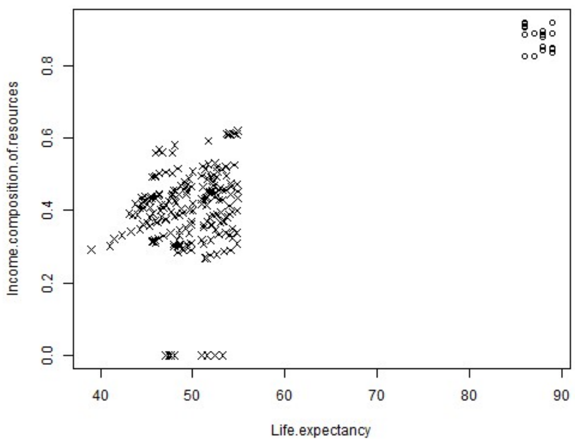

```
plot(Life.expectancy,Schooling, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="Schooling")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



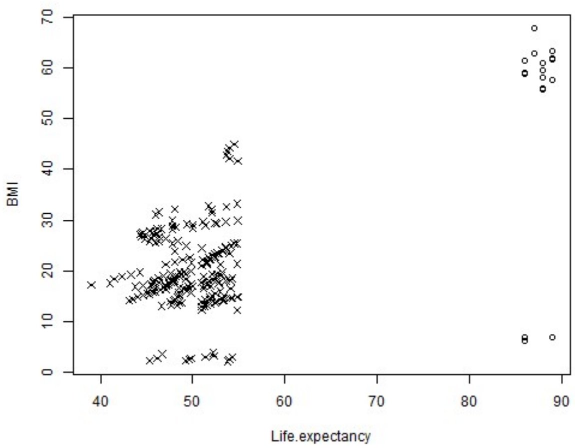

```
plot(Life.expectancy,percentage.expenditure, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="percentage.expenditure")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



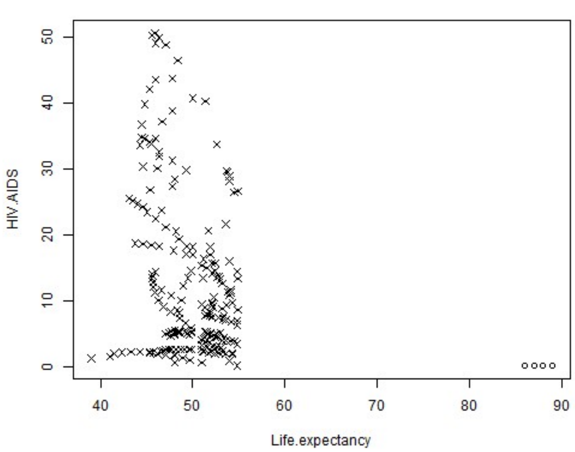

```
plot(Life.expectancy,Income.composition.of.resources, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="Income.composition.of.resources")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



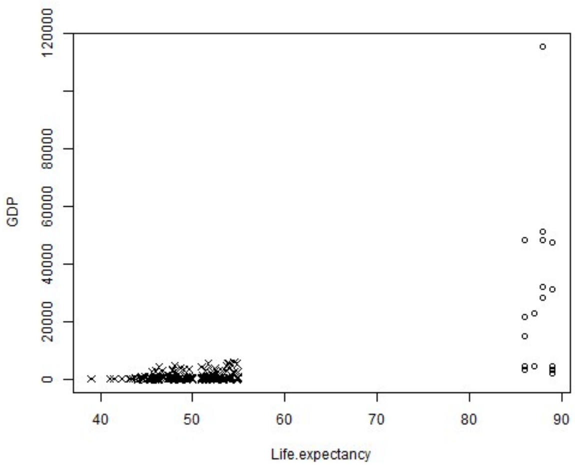

```
plot(Life.expectancy,BMI, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="BMI")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



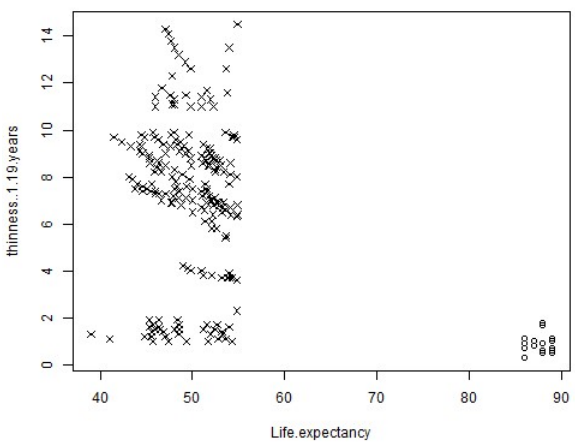

```
plot(Life.expectancy,HIV.AIDS, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="HIV.AIDS")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```



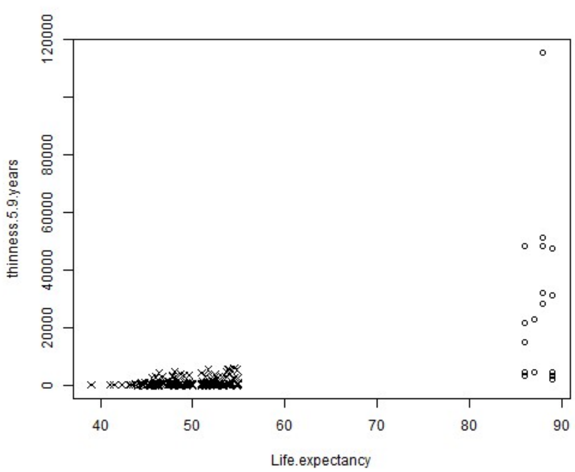

```
plot(Life.expectancy,GDP, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="GDP")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```




```r
plot(Life.expectancy,thinness..1.19.years, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="thinness..1.19.years")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```




```r
plot(Life.expectancy,GDP, pch=c(1,4)[as.numeric(Status)],xlab="Life.expectancy",ylab="thinness.5.9.years")
legend(158, 233, legend=c("Developed","Developing"), pch=c(4,1), horiz=TRUE)
```

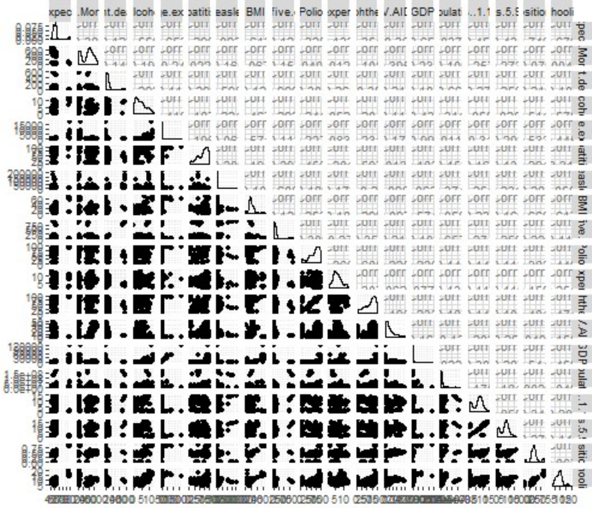



```r
# deleting the first three column as it is not the part of the dependent or independent variables
# deleting is onlt done to check the correlation of the varibles.

corr.plot <- select(new.life, -c(1,2,3))
```

```
#View(WHONew)
dim(corr.plot)
```

```
## [1] 223  19
```

```
ggpairs(as.data.frame(corr.plot[,1:length(corr.plot)]))
```



```
###################### CORRELATION TESTING ######################################
##The Pearson product-moment correlation coefficient (Pearson's correlation, for short) is a measure of the strength and direction of association tha
cor.test(new.life$Life.expectancy, new.life$Adult.Mortality)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$Adult.Mortality
## t = -6.1981, df = 221, p-value = 2.767e-09
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4913573 -0.2669404
## sample estimates:
##        cor
## -0.3848217
```

```
cor.test(new.life$Life.expectancy, new.life$infant.deaths)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$infant.deaths
## t = -1.8277, df = 221, p-value = 0.06894
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.249404288  0.009503717
## sample estimates:
##        cor
## -0.1220256
```

```
cor.test(new.life$Life.expectancy, new.life$ Alcohol)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$Alcohol
## t = 9.9525, df = 221, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4584477 0.6408564
## sample estimates:
##       cor
## 0.5563179
```

```
cor.test(new.life$Life.expectancy, new.life$percentage.expenditure)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$percentage.expenditure
## t = 12.831, df = 221, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.5710204 0.7227206
```

```
## sample estimates:
##       cor
## 0.6533814
```

```r
cor.test(new.life$Life.expectancy, new.life$Adult.Mortality)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$Adult.Mortality
## t = -6.1981, df = 221, p-value = 2.767e-09
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.4913573 -0.2669404
## sample estimates:
##        cor
## -0.3848217
```

```r
cor.test(new.life$Life.expectancy, new.life$infant.deaths)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$infant.deaths
## t = -1.8277, df = 221, p-value = 0.06894
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.249404288  0.009503717
## sample estimates:
##        cor
## -0.1220256
```

```r
cor.test(new.life$Life.expectancy, new.life$Income.composition.of.resources)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$Income.composition.of.resources
## t = 16.672, df = 221, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6818546 0.7993659
## sample estimates:
##       cor
## 0.7463716
```

```r
cor.test(new.life$Life.expectancy, new.life$Schooling)
```

```
##
##      Pearson's product-moment correlation
##
## data:  new.life$Life.expectancy and new.life$Schooling
## t = 13.619, df = 221, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.597115 0.741108
## sample estimates:
##       cor
## 0.6755009
```

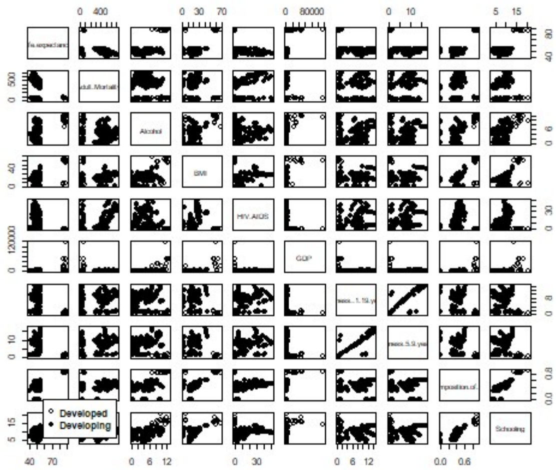
```r
###################### Plotting done for all the    ##################################

dim(new.life)
```

```
## [1] 223  22
```

```r
life <- new.life[ ,c(4,5,7,11,16,17,19,20,21,22)]
dim(life)
```

```
## [1] 223  10
```

```r
labs.diagonal <- c("Life.expectancy","Adult.Mortality","Alcohol","BMI","HIV.AIDS","GDP","thinness..1.19.years","thinness.5.9.years","Income.compositi
pairs(life, labels=labs.diagonal, pch=c(1,16)[as.numeric(Status)],font.labels = )
par(xpd=NA)
legend(-0.001,0.07,c("Developed","Developing"),pch=c(1,16), cex=0.7, text.font=2) # gives the legend points the correct symbol
```




```r
###########  TESTING OF COVARAINCE################

#Take a subset by Status
cov(subset(new.life, Status == "Developed")[,c("Life.expectancy","Adult.Mortality","Alcohol","BMI","HIV.AIDS","GDP","thinness..1.19.years","thinness.!
```

```
##                                Life.expectancy Adult.Mortality       Alcohol
## Life.expectancy                     1.55882353    -1.617647e+00  7.794118e-02
## Adult.Mortality                    -1.61764706     2.530359e+02 -1.308007e+00
## Alcohol                             0.07794118    -1.308007e+00  2.157245e+00
## BMI                                 9.28529412    -6.899902e+01  5.317873e+00
## HIV.AIDS                            0.00000000     0.000000e+00  0.000000e+00
## GDP                              6527.60661941    -8.690194e+04  1.067327e+04
## thinness..1.19.years                0.02941176    -3.085621e+00  2.469150e-01
## thinness.5.9.years                  0.05588235    -4.029085e+00  3.138791e-01
## Income.composition.of.resources    -0.01232353    -2.577451e-02  2.605892e-02
## Schooling                          -0.50294118     2.185294e+00  4.332353e-02
##                                          BMI HIV.AIDS           GDP
## Life.expectancy                  9.285294e+00        0  6.527607e+03
## Adult.Mortality                 -6.899902e+01        0 -8.690194e+04
## Alcohol                          5.317873e+00        0  1.067327e+04
## BMI                              5.355768e+02        0  2.172653e+05
## HIV.AIDS                         0.000000e+00        0  0.000000e+00
## GDP                              2.172653e+05        0  8.093754e+08
## thinness..1.19.years             8.309804e-01        0  1.539659e+03
## thinness.5.9.years               1.865784e+00        0  2.875355e+03
## Income.composition.of.resources  4.050882e-02        0  3.644499e+01
## Schooling                       -1.792618e+01        0 -1.875693e+04
##                                thinness..1.19.years thinness.5.9.years
## Life.expectancy                        2.941176e-02       5.588235e-02
## Adult.Mortality                       -3.085621e+00      -4.029085e+00
## Alcohol                                2.469150e-01       3.138791e-01
## BMI                                    8.309804e-01       1.865784e+00
## HIV.AIDS                               0.000000e+00       0.000000e+00
## GDP                                    1.539659e+03       2.875355e+03
## thinness..1.19.years                   1.516340e-01       1.889542e-01
## thinness.5.9.years                     1.889542e-01       2.413399e-01
## Income.composition.of.resources        4.460784e-03       5.859804e-03
## Schooling                             -1.135294e-01      -1.626471e-01
##                                Income.composition.of.resources     Schooling
## Life.expectancy                                   -0.012323529 -5.029412e-01
## Adult.Mortality                                   -0.025774510  2.185294e+00
## Alcohol                                            0.026058922  4.332353e-02
## BMI                                                0.040508824 -1.792618e+01
## HIV.AIDS                                           0.000000000  0.000000e+00
## GDP                                               36.444985086 -1.875693e+04
## thinness..1.19.years                               0.004460784 -1.135294e-01
## thinness.5.9.years                                 0.005859804 -1.626471e-01
## Income.composition.of.resources                    0.001053088  1.877353e-02
## Schooling                                          0.018773529  2.403824e+00
```

```r
cov(subset(new.life, Status == "Developing")[,c("Life.expectancy","Adult.Mortality","Alcohol","BMI","HIV.AIDS","GDP","thinness..1.19.years","thinness.
```

```
##                                Life.expectancy Adult.Mortality      Alcohol
## Life.expectancy                    11.54630416      -76.600947 9.132169e-01
## Adult.Mortality                   -76.60094692    43095.757628 1.259531e+01
## Alcohol                             0.91321686       12.595308 8.415275e+00
## BMI                                 2.04875012      485.505507 5.074276e+00
## HIV.AIDS                          -12.64961669      842.908663 2.213147e+00
## GDP                               850.16117969    52444.810906 1.276117e+03
## thinness..1.19.years               -0.20352224       70.253945 2.770936e+00
## thinness.5.9.years                  0.64477762       85.017396 3.311180e+00
## Income.composition.of.resources     0.06343404        8.167645 5.665478e-02
## Schooling                           0.46345409      180.574259 2.117717e+00
##                                        BMI      HIV.AIDS          GDP
## Life.expectancy                   2.048750   -12.6496167 8.501612e+02
## Adult.Mortality                 485.505507   842.9086633 5.244481e+04
```

```
## Alcohol                          5.074276    2.2131472 1.276117e+03
## BMI                             60.424370   47.2970985 4.926518e+03
## HIV.AIDS                        47.297099  157.3825528 3.888910e+03
## GDP                           4926.518423 3888.9101219 1.552012e+06
## thinness..1.19.years             2.701082    3.3166954 2.603857e+02
## thinness.5.9.years               6.162389    5.7906791 1.016913e+03
## Income.composition.of.resources  0.472489    0.6341835 7.938864e+01
## Schooling                        9.160771   14.9383752 1.283127e+03
##                               thinness..1.19.years thinness.5.9.years
## Life.expectancy                       -0.203522238       6.447776e-01
## Adult.Mortality                       70.253945481       8.501740e+01
## Alcohol                                2.770936195       3.311180e+00
## BMI                                    2.701082018       6.162389e+00
## HIV.AIDS                               3.316695361       5.790679e+00
## GDP                                  260.385745480       1.016913e+03
## thinness..1.19.years                  11.213254424       9.301809e+00
## thinness.5.9.years                     9.301808943       1.165162e+01
## Income.composition.of.resources        0.008689371       5.255204e-02
## Schooling                              1.504652080       2.223482e+00
##                               Income.composition.of.resources    Schooling
## Life.expectancy                                   0.063434041    0.4634541
## Adult.Mortality                                   8.167645170  180.5742587
## Alcohol                                           0.056654778    2.1177168
## BMI                                               0.472488981    9.1607712
## HIV.AIDS                                          0.634183517   14.9383752
## GDP                                              79.388642213 1283.1267589
## thinness..1.19.years                              0.008689371    1.5046521
## thinness.5.9.years                                0.052552044    2.2234816
## Income.composition.of.resources                   0.014380534    0.1977795
## Schooling                                         0.197779495    5.0008345
```

```r
cov(new.life[, c("Life.expectancy","Adult.Mortality","BMI","HIV.AIDS","GDP","thinness..1.19.years","thinness.5.9.years","Income.composition.of.resour
```

```
##                               Life.expectancy Adult.Mortality         BMI
## Life.expectancy                    117.819151   -8.956675e+02   83.270393
## Adult.Mortality                   -895.667533    4.597891e+04 -180.778552
## BMI                                 83.270393   -1.807786e+02  157.316086
## HIV.AIDS                           -49.037273    1.062844e+03   15.276636
## GDP                              75736.769437   -5.321624e+05 77256.195385
## thinness..1.19.years               -17.925701    2.010199e+02  -10.819543
## thinness.5.9.years                 -17.546506    2.176143e+02   -7.862754
## Income.composition.of.resources      1.415058   -2.958084e+00    1.460122
## Schooling                           22.331345   -2.983620e+00   23.576928
##                                    HIV.AIDS          GDP
## Life.expectancy               -4.903727e+01    75736.769
## Adult.Mortality                1.062844e+03  -532162.410
## BMI                            1.527664e+01    77256.195
## HIV.AIDS                       1.576927e+02   -22438.559
## GDP                           -2.243856e+04 115171785.111
## thinness..1.19.years           9.245822e+00   -11977.472
## thinness.5.9.years             1.165978e+01   -11459.741
## Income.composition.of.resources 1.084276e-01     1019.711
## Schooling                      6.060713e+00    14999.627
##                               thinness..1.19.years thinness.5.9.years
## Life.expectancy                      -1.792570e+01      -1.754651e+01
## Adult.Mortality                       2.010199e+02       2.176143e+02
## BMI                                  -1.081954e+01      -7.862754e+00
## HIV.AIDS                              9.245822e+00       1.165978e+01
## GDP                                  -1.197747e+04      -1.145974e+04
## thinness..1.19.years                  1.325472e+01       1.156777e+01
## thinness.5.9.years                    1.156777e+01       1.379923e+01
## Income.composition.of.resources      -2.165978e-01      -1.812855e-01
## Schooling                            -2.261375e+00      -1.687036e+00
##                               Income.composition.of.resources     Schooling
## Life.expectancy                                     1.4150579    22.3313453
## Adult.Mortality                                    -2.9580835    -2.9836202
## BMI                                                 1.4601217    23.5769278
## HIV.AIDS                                            0.1084276     6.0607126
## GDP                                              1019.7111530 14999.6274003
## thinness..1.19.years                               -0.2165978    -2.2613746
## thinness.5.9.years                                 -0.1812855    -1.6870363
## Income.composition.of.resources                     0.0305086     0.4613925
## Schooling                                           0.4613925     9.2760284
```

```r
###########  TESTING OF cHI SQUARE TEST   ################

##The chi-squared statistic is a single number that tells you how much difference exists between your observed counts and the counts you would expect

chisq.test(table(new.life$Life.expectancy,new.life$Adult.Mortality))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy,
## new.life$Adult.Mortality)): Chi-squared approximation may be incorrect
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  table(new.life$Life.expectancy, new.life$Adult.Mortality)
## X-squared = 14763, df = 14229, p-value = 0.0008726
```

```r
chisq.test(table(new.life$Life.expectancy,new.life$Alcohol))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$Alcohol)): Chi-
## squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$Alcohol)
## X-squared = 17432, df = 17112, p-value = 0.04232
```

```
chisq.test(table(new.life$Life.expectancy,new.life$infant.deaths))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$infant.deaths)):
## Chi-squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$infant.deaths)
## X-squared = 8927.3, df = 8277, p-value = 4.06e-07
```

```
chisq.test(table(new.life$Life.expectancy,new.life$percentage.expenditure))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy,
## new.life$percentage.expenditure)): Chi-squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$percentage.expenditure)
## X-squared = 19651, df = 19716, p-value = 0.6271
```

```
chisq.test(table(new.life$Life.expectancy,new.life$HIV.AIDS))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$HIV.AIDS)): Chi-
## squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$HIV.AIDS)
## X-squared = 14144, df = 13392, p-value = 3.192e-06
```

```
chisq.test(table(new.life$Life.expectancy,new.life$GDP))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$GDP)): Chi-
## squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$GDP)
## X-squared = 20739, df = 20646, p-value = 0.3227
```

```
chisq.test(table(new.life$Life.expectancy,new.life$Income.composition.of.resources))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy,
## new.life$Income.composition.of.resources)): Chi-squared approximation may be
## incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$Income.composition.of.resources)
## X-squared = 14813, df = 14601, p-value = 0.1077
```

```
chisq.test(table(new.life$Life.expectancy,new.life$Schooling))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$Schooling)): Chi-
## squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$Schooling)
## X-squared = 8022.6, df = 8184, p-value = 0.897
```

```
chisq.test(table(new.life$Life.expectancy,new.life$BMI))
```

```
## Warning in chisq.test(table(new.life$Life.expectancy, new.life$BMI)): Chi-
## squared approximation may be incorrect
```

```
##
##      Pearson's Chi-squared test
##
## data:  table(new.life$Life.expectancy, new.life$BMI)
## X-squared = 12559, df = 13113, p-value = 0.9997
```

```
############  t test ######################

##A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be re

#t-tests, one by one for Developed vs. developing
with(data=new.life,t.test(Life.expectancy[Status=="Developed"],Life.expectancy[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  Life.expectancy[Status == "Developed"] and Life.expectancy[Status == "Developing"]
## t = 46.966, df = 221, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  36.31386 39.49492
## sample estimates:
## mean of x mean of y
##  87.50000  49.59561
```

```
with(data=new.life,t.test(Adult.Mortality[Status=="Developed"],Adult.Mortality[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  Adult.Mortality[Status == "Developed"] and Adult.Mortality[Status == "Developing"]
## t = -5.9552, df = 221, p-value = 1.012e-08
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -388.7160 -195.4103
## sample estimates:
## mean of x mean of y
##  67.72222 359.78537
```

```
with(data=new.life,t.test(infant.deaths[Status=="Developed"],infant.deaths[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  infant.deaths[Status == "Developed"] and infant.deaths[Status == "Developing"]
## t = -2.4479, df = 221, p-value = 0.01515
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -133.85869  -14.45513
## sample estimates:
##  mean of x  mean of y
##  0.8333333 74.9902439
```

```
with(data=new.life,t.test(HIV.AIDS[Status=="Developed"],HIV.AIDS[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  HIV.AIDS[Status == "Developed"] and HIV.AIDS[Status == "Developing"]
## t = -4.4692, df = 221, p-value = 1.254e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.081855  -7.403023
## sample estimates:
## mean of x mean of y
##   0.10000  13.34244
```

```
with(data=new.life,t.test(GDP[Status=="Developed"],GDP[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  GDP[Status == "Developed"] and GDP[Status == "Developing"]
## t = 13.433, df = 221, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  22487.11 30220.07
## sample estimates:
##  mean of x  mean of y
## 27211.9370   858.3472
```

```
with(data=new.life,t.test(thinness..1.19.years[Status=="Developed"],thinness..1.19.years[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  thinness..1.19.years[Status == "Developed"] and thinness..1.19.years[Status == "Developing"]
## t = -7.9351, df = 221, p-value = 1.045e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.838962 -4.719846
## sample estimates:
## mean of x mean of y
## 0.8888889 7.1682927
```

```
with(data=new.life,t.test(thinness.5.9.years[Status=="Developed"],thinness.5.9.years[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  thinness.5.9.years[Status == "Developed"] and thinness.5.9.years[Status == "Developing"]
## t = -7.9585, df = 221, p-value = 9.024e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.012041 -4.831590
## sample estimates:
## mean of x mean of y
## 0.8611111 7.2829268
```

```
with(data=new.life,t.test(Income.composition.of.resources[Status=="Developed"],Income.composition.of.resources[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  Income.composition.of.resources[Status == "Developed"] and Income.composition.of.resources[Status == "Developing"]
## t = 16.915, df = 221, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4245732 0.5365504
## sample estimates:
## mean of x mean of y
## 0.8761667 0.3956049
```

```
with(data=new.life,t.test(Schooling[Status=="Developed"],Schooling[Status=="Developing"],var.equal=TRUE))
```

```
##
##      Two Sample t-test
##
## data:  Schooling[Status == "Developed"] and Schooling[Status == "Developing"]
## t = 14.419, df = 221, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  6.705522 8.828624
## sample estimates:
## mean of x mean of y
## 16.250000  8.482927
```