Untitled

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document 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:

df <- read.delim("C:/Rstudy/Introduction to Regression Analysis/datasets/education1970.txt",   
 header=T)  
row.names(df) = df[,1]  
df = df[,2:6]  
region <- df[,5]; region

## [1] 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 4  
## [39] 4 4 4 4 4 4 4 4 4 4 4 4

region[region == 1] = "Northeast"  
region[region == 2] = "North Central"  
region[region == 3] = "South"  
region[region == 4] = "West"  
region

## [1] "Northeast" "Northeast" "Northeast" "Northeast"   
## [5] "Northeast" "Northeast" "Northeast" "Northeast"   
## [9] "Northeast" "North Central" "North Central" "North Central"  
## [13] "North Central" "North Central" "North Central" "North Central"  
## [17] "North Central" "North Central" "North Central" "North Central"  
## [21] "North Central" "South" "South" "South"   
## [25] "South" "South" "South" "South"   
## [29] "South" "South" "South" "South"   
## [33] "South" "South" "South" "South"   
## [37] "South" "West" "West" "West"   
## [41] "West" "West" "West" "West"   
## [45] "West" "West" "West" "West"   
## [49] "West" "West"

region\_fac <- factor(region, levels = c("Northeast", "North Central", "South", "West"))  
  
df[,5] <- region\_fac;  
plot(df)  
cor(df[,1:4])

## Y X1 X2 X3  
## Y 1.0000000 0.6599045 0.3213438 0.2293063  
## X1 0.6599045 1.0000000 -0.1592206 0.6516760  
## X2 0.3213438 -0.1592206 1.0000000 -0.1367213  
## X3 0.2293063 0.6516760 -0.1367213 1.0000000

library(HH)

## 필요한 패키지를 로딩중입니다: lattice

## 필요한 패키지를 로딩중입니다: grid

## 필요한 패키지를 로딩중입니다: latticeExtra

## 필요한 패키지를 로딩중입니다: multcomp

## 필요한 패키지를 로딩중입니다: mvtnorm

## 필요한 패키지를 로딩중입니다: survival

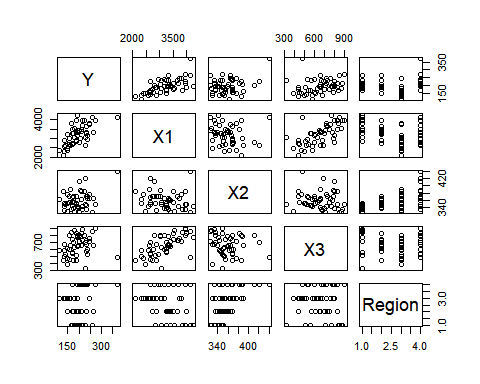
## 필요한 패키지를 로딩중입니다: TH.data

## 필요한 패키지를 로딩중입니다: MASS

##   
## 다음의 패키지를 부착합니다: 'TH.data'

## The following object is masked from 'package:MASS':  
##   
## geyser

## 필요한 패키지를 로딩중입니다: gridExtra



vif(df[2:4])

## X1 X2 X3   
## 1.753440 1.028002 1.741542

result\_a <- lm(Y~X1+X2+X3, data = df[,1:4])  
summary(result\_a)

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3, data = df[, 1:4])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -60.84 -15.67 -0.87 15.73 51.12   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.892e+02 6.617e+01 -4.370 7.01e-05 \*\*\*  
## X1 8.089e-02 9.474e-03 8.538 4.83e-11 \*\*\*  
## X2 8.184e-01 1.616e-01 5.064 7.10e-06 \*\*\*  
## X3 -1.038e-01 3.506e-02 -2.960 0.00486 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 26.97 on 46 degrees of freedom  
## Multiple R-squared: 0.6825, Adjusted R-squared: 0.6618   
## F-statistic: 32.96 on 3 and 46 DF, p-value: 1.592e-11

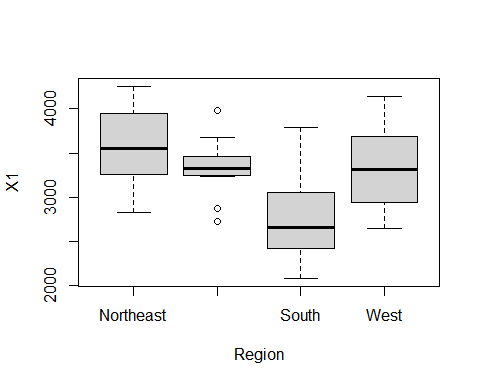
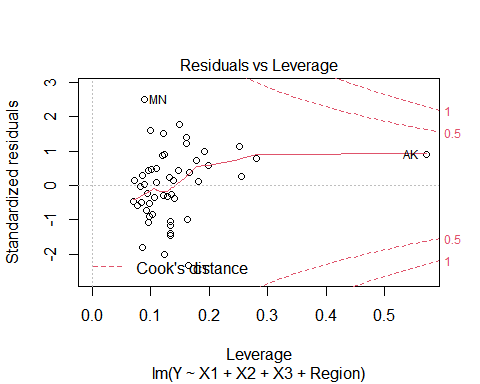
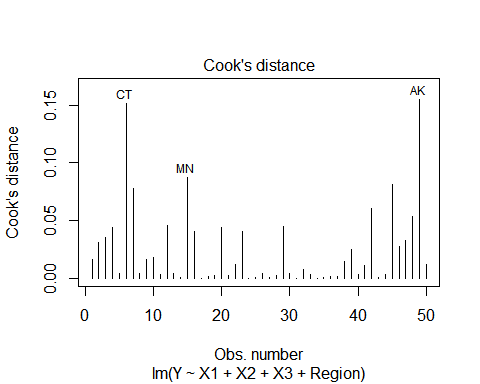
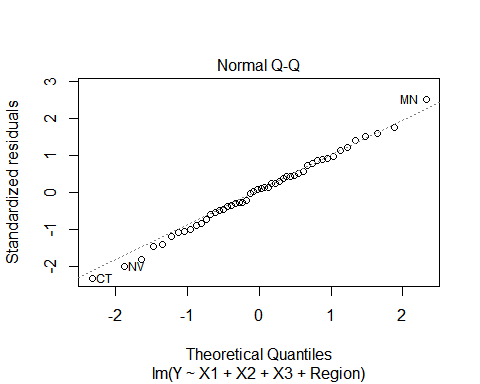
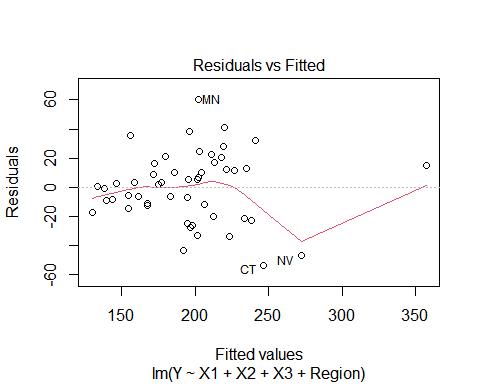
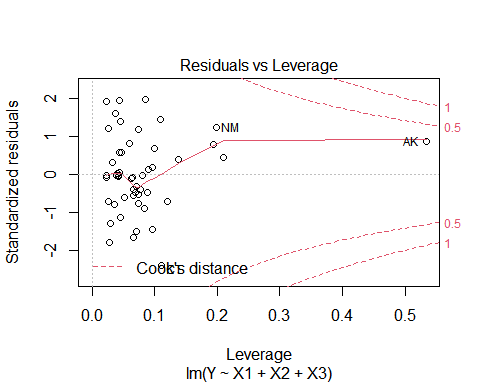
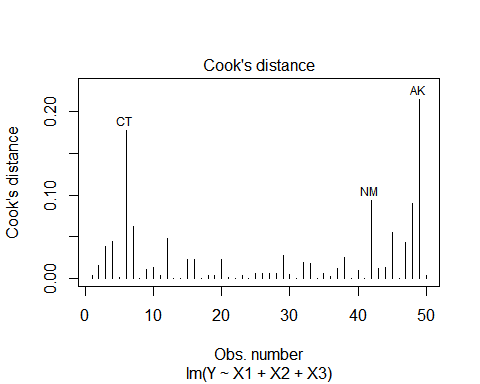
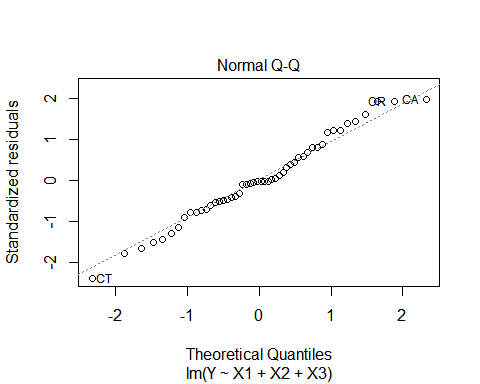
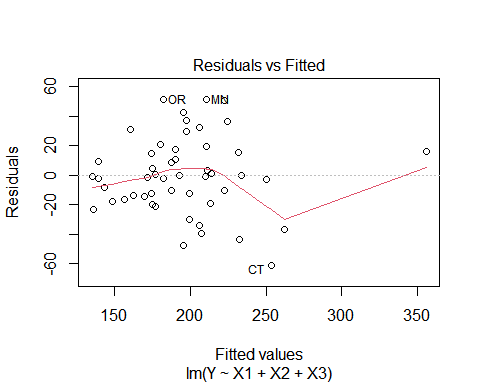
result\_e <- lm(Y~X1+X2+X3+Region, data = df); summary(result\_e)

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3 + Region, data = df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -53.518 -14.030 2.348 14.433 60.176   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -177.14303 72.35159 -2.448 0.018505 \*   
## X1 0.07878 0.01047 7.525 2.25e-09 \*\*\*  
## X2 0.54048 0.18471 2.926 0.005463 \*\*   
## X3 -0.12392 0.03439 -3.604 0.000809 \*\*\*  
## RegionNorth Central 0.78393 11.75901 0.067 0.947156   
## RegionSouth -0.56818 13.00634 -0.044 0.965358   
## RegionWest 29.79412 13.63529 2.185 0.034380 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 25.26 on 43 degrees of freedom  
## Multiple R-squared: 0.7397, Adjusted R-squared: 0.7033   
## F-statistic: 20.36 on 6 and 43 DF, p-value: 4.057e-11

## Including Plots

You can also embed plots, for example:

##   
## Call:  
## lm(formula = Y ~ X1 + X2 + X3, data = df[, 1:4])  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -60.84 -15.67 -0.87 15.73 51.12   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.892e+02 6.617e+01 -4.370 7.01e-05 \*\*\*  
## X1 8.089e-02 9.474e-03 8.538 4.83e-11 \*\*\*  
## X2 8.184e-01 1.616e-01 5.064 7.10e-06 \*\*\*  
## X3 -1.038e-01 3.506e-02 -2.960 0.00486 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 26.97 on 46 degrees of freedom  
## Multiple R-squared: 0.6825, Adjusted R-squared: 0.6618   
## F-statistic: 32.96 on 3 and 46 DF, p-value: 1.592e-11



Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.