**R Code for Examples in the book**



***“Statistics: The Art and Science of Learning from Data”***

**by Agresti, Franklin and Klingenberg, 5th edition**

**Chapter 3**

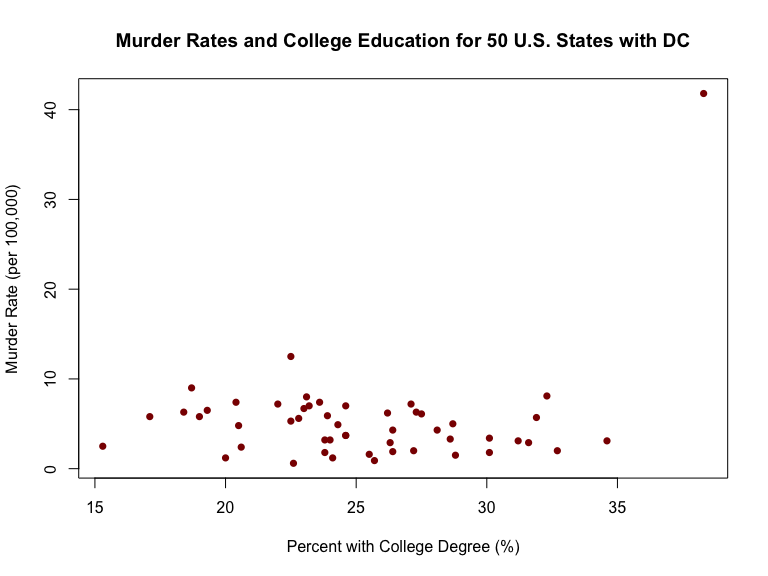
**Example 14: Education and Murder – Influential Outliers**

## Reading in the data

crime <- read.csv(file='https://raw.githubusercontent.com/artofstat/data/master/Chapter3/us\_statewide\_crime.csv')  
attach(crime) # so we can refer to variable names

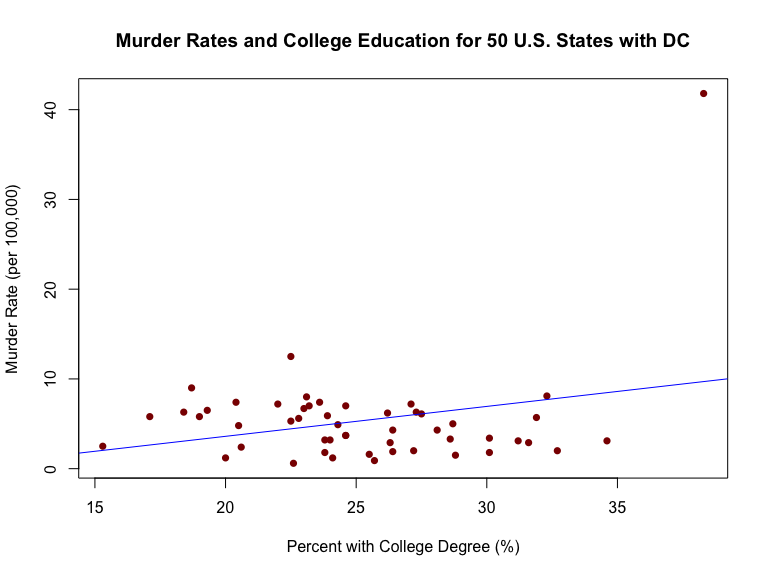
## Basic scatterplot

plot(x = college, y = murder.rate, pch = 16, col = 'darkred',  
 main = 'Murder Rates and College Education for 50 U.S. States with DC',  
 xlab = 'Percent with College Degree (%)',   
 ylab = 'Murder Rate (per 100,000)')



## Fitting in regression model

lin.reg <- lm(murder.rate ~ college)  
plot(x = college, y = murder.rate, pch = 16, col = 'darkred',  
 main = 'Murder Rates and College Education for 50 U.S. States with DC',  
 xlab = 'Percent with College Degree (%)',   
 ylab = 'Murder Rate (per 100,000)')  
abline(lin.reg, col = 'blue')



lin.reg

##   
## Call:  
## lm(formula = murder.rate ~ college)  
##   
## Coefficients:  
## (Intercept) college   
## -3.0581 0.3331

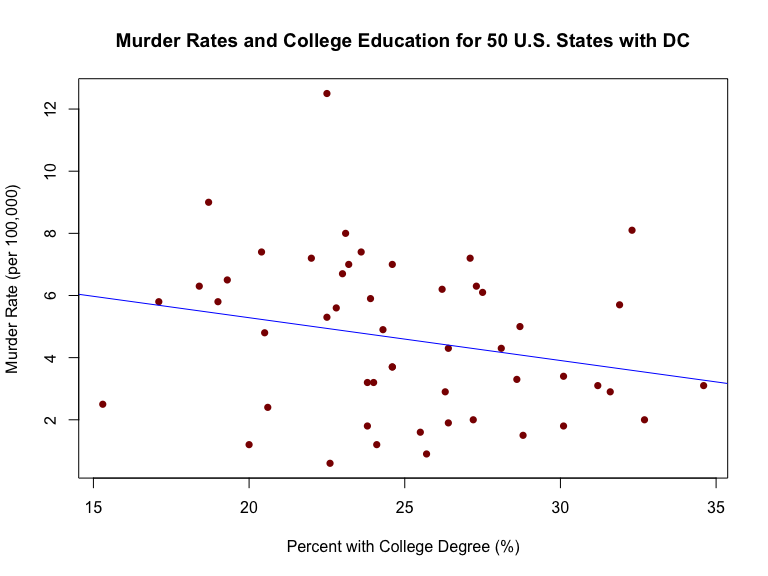
detach(crime)

## Excluding the observation for DC

crime <- subset(crime, State != 'District of Columbia')

## Fitting in new regression model

attach(crime)  
lin.reg <- lm(murder.rate ~ college)  
plot(x = college, y = murder.rate, pch = 16, col = 'darkred',  
 main = 'Murder Rates and College Education for 50 U.S. States with DC',  
 xlab = 'Percent with College Degree (%)',   
 ylab = 'Murder Rate (per 100,000)')  
abline(lin.reg, col = 'blue')



lin.reg

##   
## Call:  
## lm(formula = murder.rate ~ college)  
##   
## Coefficients:  
## (Intercept) college   
## 8.0416 -0.1379