**R Code for Examples in the book**



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

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

**Chapter 2**

**Example 17: Pollution Outliers – z-Scores**

## Read in CO2 pollution values:

eu\_co2 <- read.csv(file='http://www.artofstats.com/data/chapter2/EU\_CO2.csv')  
attach(eu\_co2) # so we can refer to variable names

## To find the z-score for the CO2 value of Luxembourg

z\_score\_Luxembourg <- (21.4 - mean(CO2)) / sd(CO2)  
z\_score\_Luxembourg

## [1] 3.749518

## To find the z-score for the CO2 value of the United States

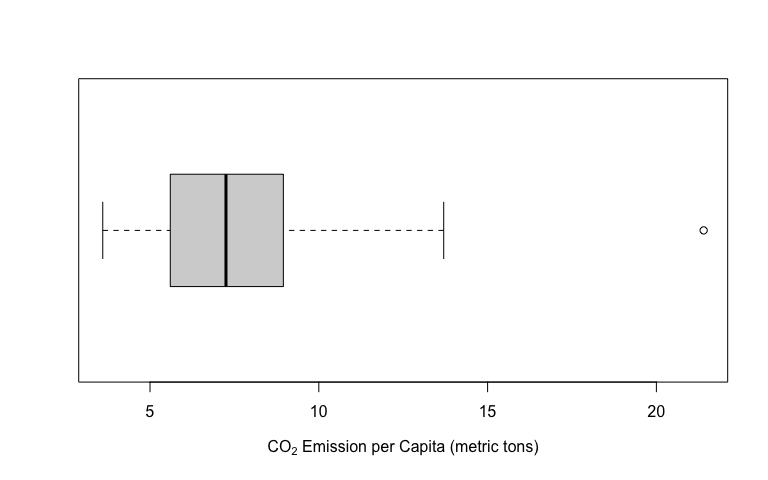
z\_score\_US <- (16.9 - mean(CO2)) / sd(CO2)  
z\_score\_US

## [1] 2.502977

## 

## Basic Box Plot

boxplot(CO2, horizontal = TRUE,   
 xlab = expression('CO'[2]\*' Emission per Capita (metric tons)'))



## 

## A better-looking box plot can be obtained with the ggplot2 library. To install it, type install.packages('ggplot2').

library(ggplot2)  
ggplot(data.frame(CO2), aes(x = '', y = CO2)) +   
 geom\_boxplot(width = 0.3, fill = 'tan') +  
 coord\_flip() +  
 labs(x = '', y = expression('CO'[2]\*' Emission per Capita (metric tons)'),   
 title = 'EU Air Pollution Data') +  
 theme\_classic() +   
 scale\_y\_continuous(limits = c(0,24),   
 breaks = seq(0,22,2),   
 expand = c(0,0)) +   
 theme(axis.line.y=element\_blank(),  
 axis.text.y=element\_blank(),  
 axis.ticks.y=element\_blank(),  
 axis.title.y=element\_blank()  
 )

