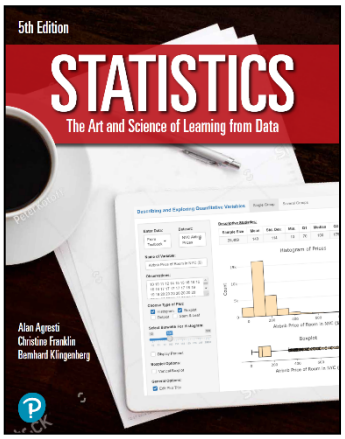


R Code for Examples in the book
"Statistics: The Art and Science of Learning from Data"
by Agresti, Franklin and Klingenberg, 5th edition

Chapter 2

Example 17: Pollution Outliers – z-Scores

Read in CO2 pollution values:

```
euCO2 <-
read.csv(file='https://raw.githubusercontent.com/artofstat/data/master/Chapter2/EU_CO2.csv')
attach(euCO2) # so we can refer to variable names
```

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

```
zScoreLuxembourg <- (21.4 - mean(CO2)) / sd(CO2)
zScoreLuxembourg

## [1] 3.749518
```

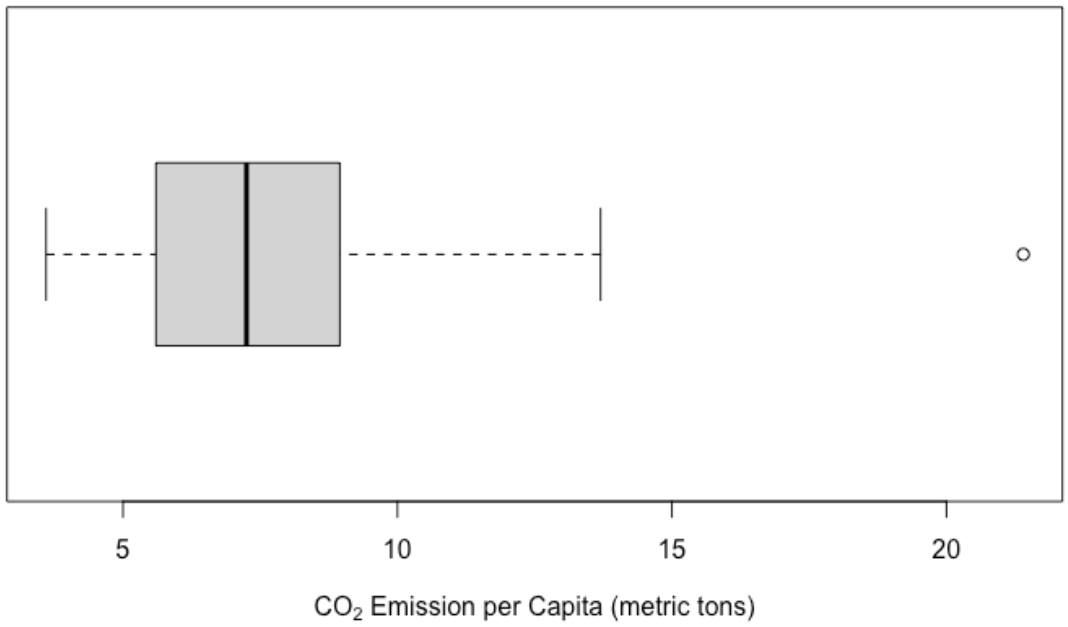
To find the z-score for the CO2 value of the United States

```
zScoreUS <- (16.9 - mean(CO2)) / sd(CO2)
zScoreUS

## [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()
        )
```