**R Code for Examples in the book**



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

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

**Chapter 8**

**Example 13: Carbon Footprint – Confidence Interval for the Correlation**

## Reading in the data:

data <- read.csv(file='https://raw.githubusercontent.com/artofstat/data/master/Chapter7/carbon\_footprint\_sandwich.csv')  
attach(data) # so we can refer to variable names

## To compute the correlation coefficient between carbon footprint and energy content

cor(EnergyContent..kCal., Carbon.footprint..g.CO2.eq..)

## [1] 0.6208991

## To obtain a bootstrap sample of the sandwiches

sample(Sandwich, replace = TRUE)

## [1] "Prawn, Mayo" "Tuna, Cucumber"   
## [3] "Chicken, Sweetcorn" "Chicken, Bacon"   
## [5] "Ham, Salad" "Ham, Mustard"   
## [7] "Sausage, Brown Sauce" "Tuna, Cucumber"   
## [9] "Prawn, Mayo" "Ham, Cheese"   
## [11] "Prawn, Mayo" "Cheese, Tomato"   
## [13] "Tuna, Cucumber" "Ham, Egg"   
## [15] "Prawn, Mayo" "Chicken, Bacon"   
## [17] "Ham, Mayo" "Ham, Salad"   
## [19] "Cheese, Mayo" "Ham, Cheese"   
## [21] "Breakfast" "Bacon, Lettuce, Tomato (BLT)"  
## [23] "Ham, Mayo" "Ham, Egg"

## 

## To obtain a bootstrap sample of the rows of the dataframe, you can use data[sample(seq\_len(nrow(data)), replace = TRUE), ]

## To generate 10,000 bootstrap samples and find the correlation

bootcorr <- c() # initializing  
for (i in 1:10000) {  
 bootsample <- data[sample(seq\_len(nrow(data)), replace = TRUE), ]  
 bootcorr[i] <- cor(bootsample$EnergyContent..kCal., bootsample$Carbon.footprint..g.CO2.eq..)  
}

## To obtain summary of the correlation coefficients from the bootstrap samples

quantile(bootcorr, c(0.025, 0.975))

## 2.5% 97.5%   
## 0.2610255 0.8379493