# Bivariate and Multivariate Graphical Data Analysis

## 1. Bivariate analysis

#### Covariance Code

## Example  
# ---  
# Question: Find the covariance of eruption duration and waiting time in the data set faithful   
# ---  
#   
  
# Printing out the the first 6 rows of the dataset  
# ---  
#   
head(faithful)

## eruptions waiting  
## 1 3.600 79  
## 2 1.800 54  
## 3 3.333 74  
## 4 2.283 62  
## 5 4.533 85  
## 6 2.883 55

# Assigning the eruptions column to the variable eruptions  
# ---  
#   
eruptions <- faithful$eruptions  
  
# Assigning the waiting column to the variable waiting  
# ---  
#   
waiting<- faithful$waiting  
  
# Using the cov() function to determine the covariance  
# ---  
#  
cov(eruptions, waiting)

## [1] 13.97781

## Challenge  
# ---  
# Question: Find out the covariance of Bwt and Hwt in the cats dataset  
  
#   
  
# Previewing the cats dataset  
# ---  
#   
library(MASS)  
head(cats)

## Sex Bwt Hwt  
## 1 F 2.0 7.0  
## 2 F 2.0 7.4  
## 3 F 2.0 9.5  
## 4 F 2.1 7.2  
## 5 F 2.1 7.3  
## 6 F 2.1 7.6

bwt <- cats$Bwt  
  
hwt <- cats$Hwt  
  
cat.cov <- cov(bwt,hwt)  
cat.cov

## [1] 0.9501127

#### Correlation Coefficient

## Example   
# ---  
# Question: Find the correlation coefficient of eruption duration and waiting time in the faithful dataset  
# ---  
  
# Assigning the eruptions column to the variable eruptions  
# ---  
#   
eruptions <- faithful$eruptions  
  
# Assigning the waiting column to the variable waiting  
# ---  
#  
waiting<- faithful$waiting  
  
# Using the cor() function to determine the covariance  
# ---  
#  
cor(eruptions, waiting)

## [1] 0.9008112

## Challenge   
# ---  
# Question: Find out the covariance of Bwt and Hwt in the cats data set below:  
# ---  
#   
  
# Previewing the cats dataset by first importing the Mass library   
# then displaying the first 6 records of this database  
library(MASS)  
head(cats)

## Sex Bwt Hwt  
## 1 F 2.0 7.0  
## 2 F 2.0 7.4  
## 3 F 2.0 9.5  
## 4 F 2.1 7.2  
## 5 F 2.1 7.3  
## 6 F 2.1 7.6

bwt <- cats$Bwt  
hwt <- cats$Hwt  
cor(bwt,hwt)

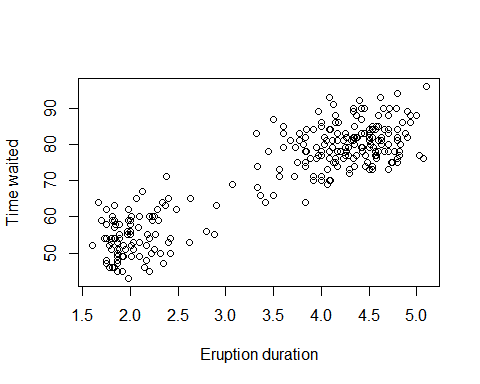
## [1] 0.8041274

## Challenge  
# ---  
# Question: Create a correlation matrix in R using the corr() function  
# ---  
# Hint: http://bit.ly/RDocumentationCorrMatrix  
# ---  
  
# Dataset url = http://bit.ly/HousingDatainR  
# ---  
#library(data.table)  
#house <- fread("http://bit.ly/HousingDatainR")  
#head(house)

## 2. Graphical Techniques

#### Scatterplot Code

## Example   
# ---  
# Question: Create a scatter plot of the eruption durations and waiting intervals from the faithful dataset  
# ---  
# OUR CODE GOES BELOW   
#   
  
# Assigning the eruptions column to the variable eruptions  
# ---  
#   
eruptions <- faithful$eruptions  
  
# Assigning the waiting column to the variable waiting  
# ---  
#  
waiting <- faithful$waiting  
  
# Creating the scatter plot using eruptions and waiting  
# ---  
#   
plot(eruptions, waiting, xlab="Eruption duration", ylab="Time waited")



# Challenge   
# ---  
# Question: Using the cats dataset, create a scatter plot of the Bwt and Hwt variables.   
# Does it reveal any relationship between these variables?  
# ---  
  
#   
  
# Previewing the cats dataset  
# ---  
#   
library(MASS)  
head(cats)

## Sex Bwt Hwt  
## 1 F 2.0 7.0  
## 2 F 2.0 7.4  
## 3 F 2.0 9.5  
## 4 F 2.1 7.2  
## 5 F 2.1 7.3  
## 6 F 2.1 7.6

bwt <- cats$Bwt  
hwt <- cats$Hwt  
plot(bwt,hwt, xlab="Bwt", ylab="Hwt")

