Week 1

#create vector of numbers  
vec<-c(8, 3, 8, 7, 15, 9, 12, 4, 9, 10, 5, 1)  
  
#create matrix with 5 rows and 2 columns from the values in vec  
mat = matrix(vec, nrow = 5, ncol = 2)

## Warning in matrix(vec, nrow = 5, ncol = 2): data length [12] is not a sub-  
## multiple or multiple of the number of rows [5]

#set the names of rows in mat  
row.names(mat)<-c("r1", "r2", "r3", "r4", "r5")  
  
#add a column to the matrix  
mat<-cbind(mat, c(1,3,5,7,9))  
  
#print the matrix sorted by first column  
mat[order(mat[,1],decreasing=FALSE),]

## [,1] [,2] [,3]  
## r2 3 12 3  
## r4 7 9 7  
## r1 8 9 1  
## r3 8 4 5  
## r5 15 10 9

#load the library UsingR  
library(UsingR)

## Loading required package: MASS

## Loading required package: HistData

## Loading required package: Hmisc

## Loading required package: lattice

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##   
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:base':  
##   
## format.pval, units

##   
## Attaching package: 'UsingR'

## The following object is masked from 'package:survival':  
##   
## cancer

#save the first and second column of homedata in vectors  
y1970price = homedata$y1970  
y2000price = homedata$y2000  
  
#find min and max price in 2000 year  
minPrice = min(y2000price)  
maxPrice = max(y2000price)  
  
#return the price of the house that had min price in 2000 in 1970  
y1970price[which(y2000price == minPrice)]

## [1] 10000

#return the price of the house that had max price in 2000 in 1970  
y1970price[which(y2000price == maxPrice)]

## [1] 198900

#return the five most expensive houses in 2000  
tail(sort(y2000price), 5)

## [1] 988900 1042000 1085000 1093500 1182800

#the number of houses more expensive than 750,000 in year 2000  
sum(y2000price > 750000)

## [1] 19

#find the prices of those with price more than 750,000  
mostExpensive = y2000price[which(y2000price > 750000)]  
#or  
mostExpensive = y2000price[y2000price > 750000]  
  
#The mean value of mostExpensive  
mean(mostExpensive)

## [1] 913300

#the price of the houses which price has decreased from 1970 to 2000  
y2000price[which(y1970price > y2000price)]

## [1] 7400

homedata[which(y1970price > y2000price), 0:2]

## y1970 y2000  
## 2903 10000 7400

#vector which contains with how much percent has the price increased  
x = (y2000price - y1970price)/y1970price\*100  
  
#the increase of the 10 houses with biggest increase in price  
head(sort(x, decreasing = TRUE), 10)

## [1] Inf 831.4286 745.8333 587.0455 571.8266 566.7707 563.3721 546.3368  
## [9] 545.6494 543.7164

#или  
head(homedata[order(homedata$y1970/homedata$y2000),], 10)

## y1970 y2000  
## 2048 0 432600  
## 3522 17500 163000  
## 334 19200 162400  
## 2435 44000 302300  
## 220 161500 1085000  
## 2536 64100 427400  
## 2533 86000 570500  
## 3523 105100 679300  
## 341 79300 512000  
## 274 93100 599300

#load the library  
library(MASS)  
  
#find number of males in dataframe na.rm remove the N/A values from the count  
sum(survey$Sex == 'Male', na.rm = TRUE)

## [1] 118

#number of males who smoke  
sum(survey$Sex == 'Male' & survey$Smoke != 'Never', na.rm = TRUE)

## [1] 28

#mean height of males  
mean(survey$Height, na.rm = TRUE)

## [1] 172.3809

#height and gender of 6th youngest students  
head(survey[order(survey$Age),c('Height', 'Sex')])

## Height Sex  
## 81 NA Male  
## 3 NA Male  
## 68 NA Female  
## 234 160.00 Female  
## 107 172.00 Female  
## 197 170.18 Female