#R code: Discussion 10.

#The following is NOT required for your homework

#Optional Topic: More on exploratory data analysis

?summary
?hist
?plot
?par
?points
?legend

#Example: Grocery Retailer: Problem 6.9
Data = read.table("CH06PR09.txt")
names(Data) = c("Hours", "Cases", "Costs", "Holiday")

#get summary statistics of the data
summary(Data)

#allow more than 1 plot per same page
par(mfrow=c(1,3))  #split into 1 row, 3 columns

#plot histograms of the predictors and response
hist(Data$Cases, xlab="Cases", main="Histogram of Cases")
hist(Data$Costs, nclass = 10)
  #specify the number of bars
hist(Data$Hours, breaks = seq(from=3800, to=5200, by=200))
  #specify the cluster boundaries

#restore back to original: single graph per screen
par(mfrow=c(1,1))

#Drop to the end to see the final command..

par(mfrow=c(2,2))  #split into 2 rows, 2 columns

#plot Hours (Y) vs. Cases (X1), with open circles
plot(Data$Cases, Data$Hours,
     xlab="Cases", ylab="Hours", main="Scatterplot of Hours vs. Cases")

#plot Hours (Y) vs. Cases (X1), with filled circles
plot(Data$Cases, Data$Hours,
     xlab="Cases", ylab="Hours", main="Scatterplot of Hours vs. Cases", pch=19)

#plot Hours (Y) vs. Cases (X1), with different characters for Holiday (X3)
#create empty plot, with labels
#add points by 'triangle' => pch=17, if Holiday=1
#add points by 'circle' => pch=19, if Holiday=0
plot(Data$Cases, Data$Hours,
     xlab="Cases", ylab="Hours", main="Scatterplot of Hours vs. Cases", type="n")
points(Data$Cases[Data$Holiday == 1], Data$Hours[Data$Holiday == 1], pch=17)
points(Data$Cases[Data$Holiday == 0], Data$Hours[Data$Holiday == 0], pch=19)
#color the points
plot(Data$Cases, Data$Hours,
     xlab="Cases", ylab="Hours", main="Scatterplot of Hours vs. Cases",
     type="n")
points(Data$Cases[Data$Holiday == 1], Data$Hours[Data$Holiday == 1],
       pch=17, col="red")
points(Data$Cases[Data$Holiday == 0], Data$Hours[Data$Holiday == 0],
       pch=19, col="blue")

#add the legend, use same parameters as in the plot above
legend("topleft", legend=c("Holiday","No Holiday"),
       pch=c(17,19), col=c("red","blue"))
#First argument is position of legend: possible choices are:
#"bottomright", "bottom", "bottomleft", "left", "topleft", "top",
#"topright", "right" and "center", and
# locator(1), a manual positioning

#Final command:
par(mfrow=c(1,1))
plot(Data$Cases, Data$Hours,
     xlab="Cases", ylab="Hours", main="Scatterplot of Hours vs. Cases",
     type="n")
points(Data$Cases[Data$Holiday == 1], Data$Hours[Data$Holiday == 1],
       pch=17, col="red")
points(Data$Cases[Data$Holiday == 0], Data$Hours[Data$Holiday == 0],
       pch=19, col="blue")
legend(locator(1), legend=c("Holiday","No Holiday"),
       pch=c(17,19), col=c("red","blue"))

#Remember, locator(1) waits for you to click on the screen to add the legend where you want

#Plot X1 vs. X2
plot(Data$Cases, Data$Costs,
     xlab="Cases", ylab="Costs", main="Scatterplot of Cases vs. Costs",
     type="n")
points(Data$Cases[Data$Holiday == 1], Data$Costs[Data$Holiday == 1],
       pch=17, col="red")
points(Data$Cases[Data$Holiday == 0], Data$Costs[Data$Holiday == 0],
       pch=19, col="black")

#add the legend, use same parameters as in the plot above
legend(locator(1), legend=c("Holiday","No Holiday"),
       pch=c(17,19), col=c("red","black"))