HW2_Math564

V. V. Dhanvanthar M. [A20543395]

2024-09-11

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com (http://rmarkdown.rstudio.com).

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Question 2

```
x <- c(1, 2, 4, 6, 7, 8, 10)
y <- c(10, 14, 12, 13, 15, 12, 13)
#combine vectors into data frame
df <- data.frame(x=x, y=y)
# print the data frame
print(df)
```

```
##    x  y
## 1  1 10
## 2  2 14
## 3  4 12
## 4  6 13
## 5  7 15
## 6  8 12
## 7 10 13
```

Question 4

```
#a
#manual computing

Sxx = sum(x*x)-length(x)*mean(x)^2
Sxy = sum(x*y)-length(x)*mean(x)*mean(y)
Syy = sum(y*y)-length(x)*mean(y)*2

#print Sxx, Sxy, Syy
print(paste("Sxx = ", Sxx))
```

```
## [1] "Sxx =  63.7142857142857"
```

```
print(paste("Sxy = ", Sxy))
```

```
## [1] "Sxy =  11.8571428571429"
```

```
print(paste("Syy = ", Syy))
```

```
## [1] "Syy =  969"
```

```
beta1hat = Sxy/Sxx
beta0hat = mean(y)-beta1hat*mean(x)
#print b1hat, b0hat
print(paste("beta0hat =", beta0hat))
```

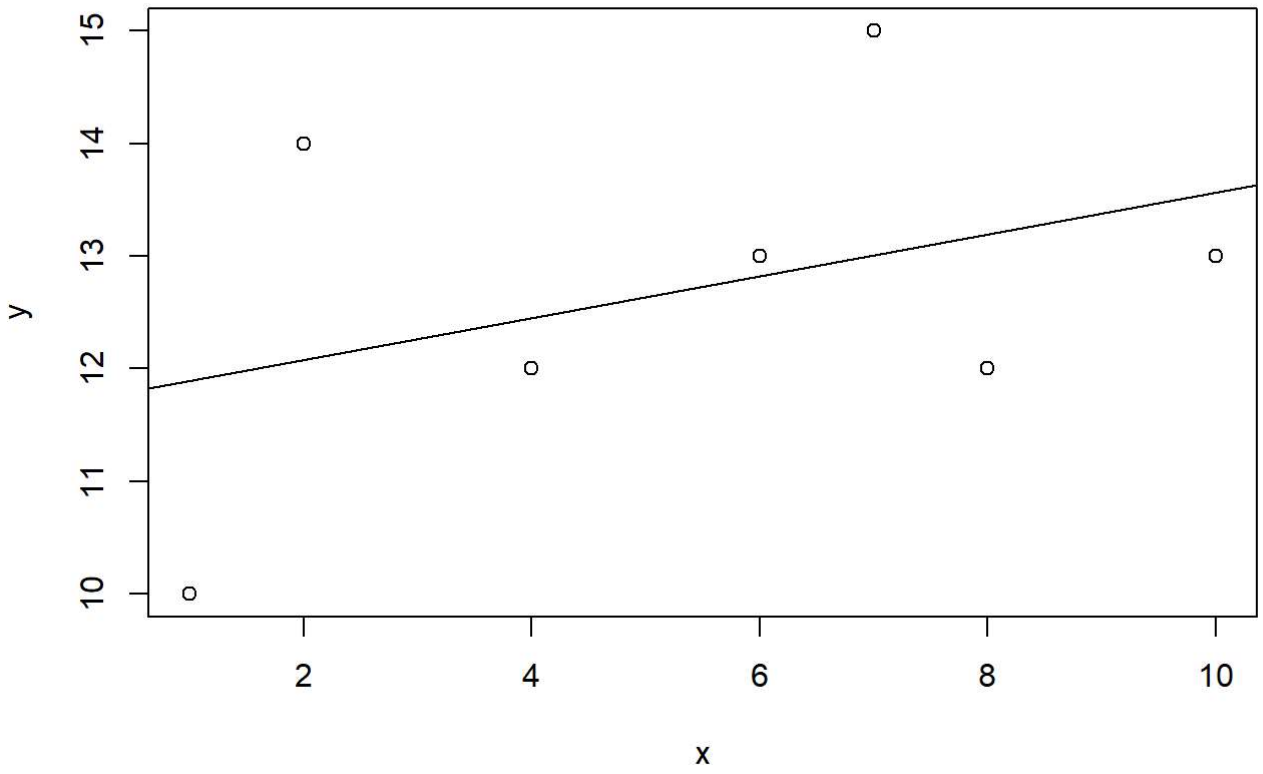
```
## [1] "beta0hat = 11.7040358744395"
```

```
print(paste("beta1hat = ", beta1hat))
```

```
## [1] "beta1hat =  0.186098654708521"
```

```
#b

#abline(a,b) adds a line with y-intercept as a and slope as b
plot(x,y)
abline(beta0hat, beta1hat)
```



Question 5

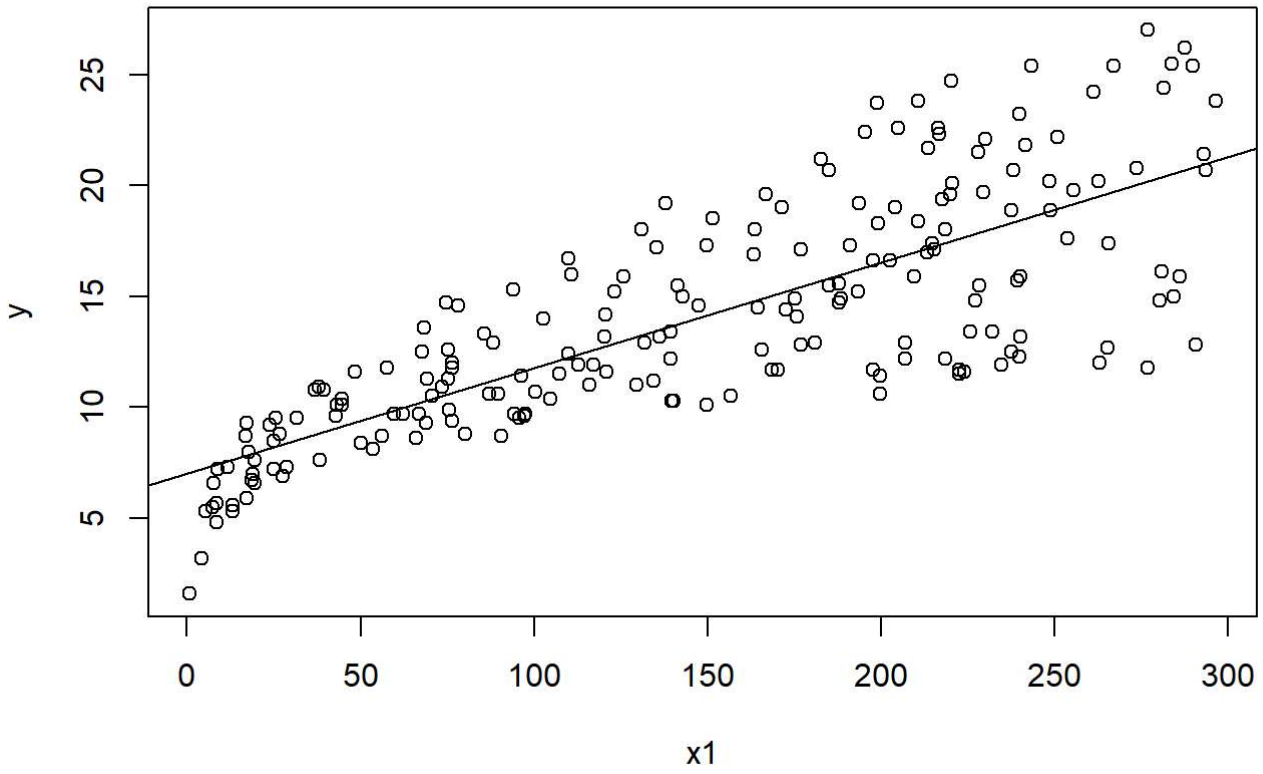
```
#(ii)
df <- read.csv("Advertising.CSV", header = TRUE, sep =",")
#re-assigning the feature labels
x1 <- df$TV
x2 <- df$radio
x3 <- df$newspaper
y <- df$sales

#using builtin regression function

model1<- lm(y ~ x1)
options(digits = 9)
summary(model1)
```

```
##
## Call:
## lm(formula = y ~ x1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -8.385982 -1.954522 -0.191265  2.067109  7.212369
##
## Coefficients:
##               Estimate Std. Error t value   Pr(>|t|)
## (Intercept) 7.03259355 0.45784294 15.3603 < 2.22e-16 ***
## x1          0.04753664 0.00269061 17.6676 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.25866 on 198 degrees of freedom
## Multiple R-squared:  0.611875,   Adjusted R-squared:  0.609915
## F-statistic: 312.145 on 1 and 198 DF,  p-value: < 2.22e-16
```

```
plot(x1,y)
abline(model1)
```


y
5
10
15
20
25
0
50
100
150
200
250
300
x1

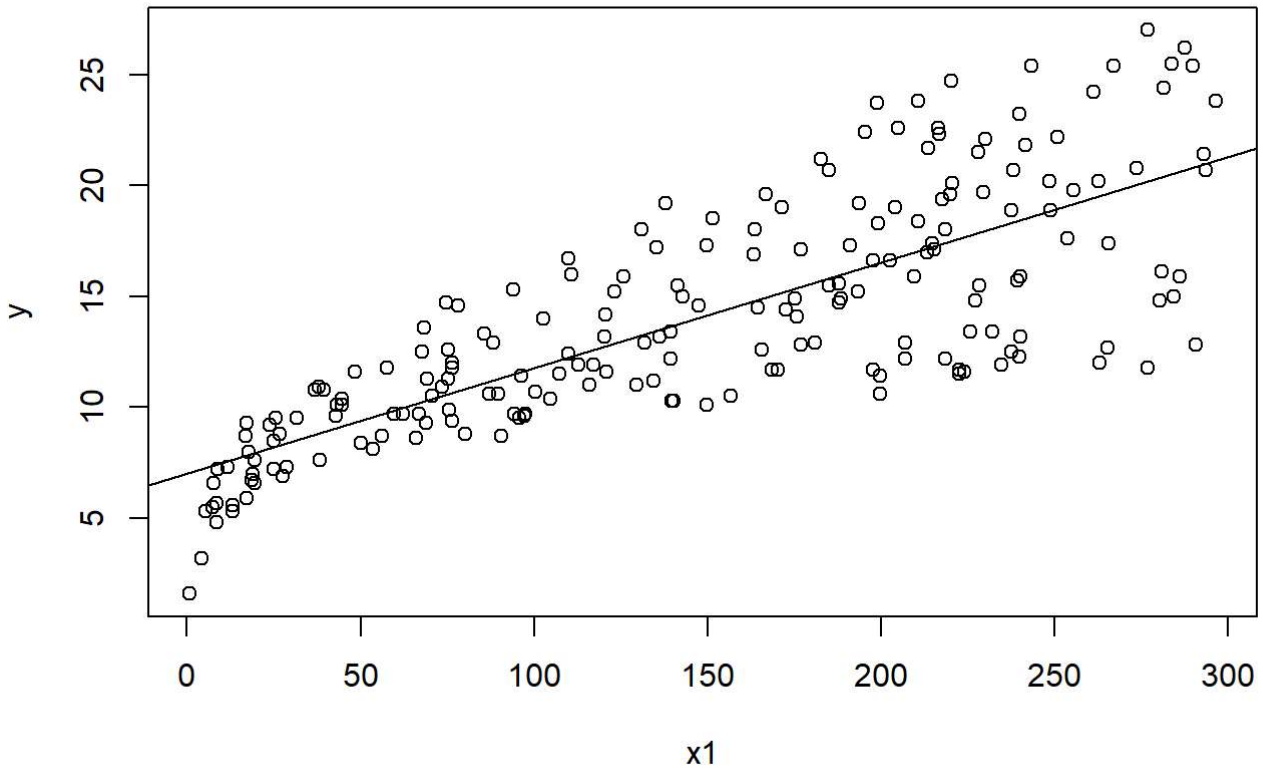
```
#(iii)
f <- read.csv("Advertising.CSV", header = TRUE, sep =",")
#re-assigning the feature labels
x1 <- df$TV
x2 <- df$radio
x3 <- df$newspaper
y <- df$sales

#using builtin regression function
#for y onto x1

model1<- lm(y ~ x1)
options(digits = 9)
summary(model1)
```

```
##
## Call:
## lm(formula = y ~ x1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -8.385982 -1.954522 -0.191265  2.067109  7.212369
##
## Coefficients:
##               Estimate Std. Error t value   Pr(>|t|)
## (Intercept) 7.03259355 0.45784294 15.3603 < 2.22e-16 ***
## x1          0.04753664 0.00269061 17.6676 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.25866 on 198 degrees of freedom
## Multiple R-squared:  0.611875,   Adjusted R-squared:  0.609915
## F-statistic: 312.145 on 1 and 198 DF,  p-value: < 2.22e-16
```

```
plot(x1,y)
abline(model1)
```

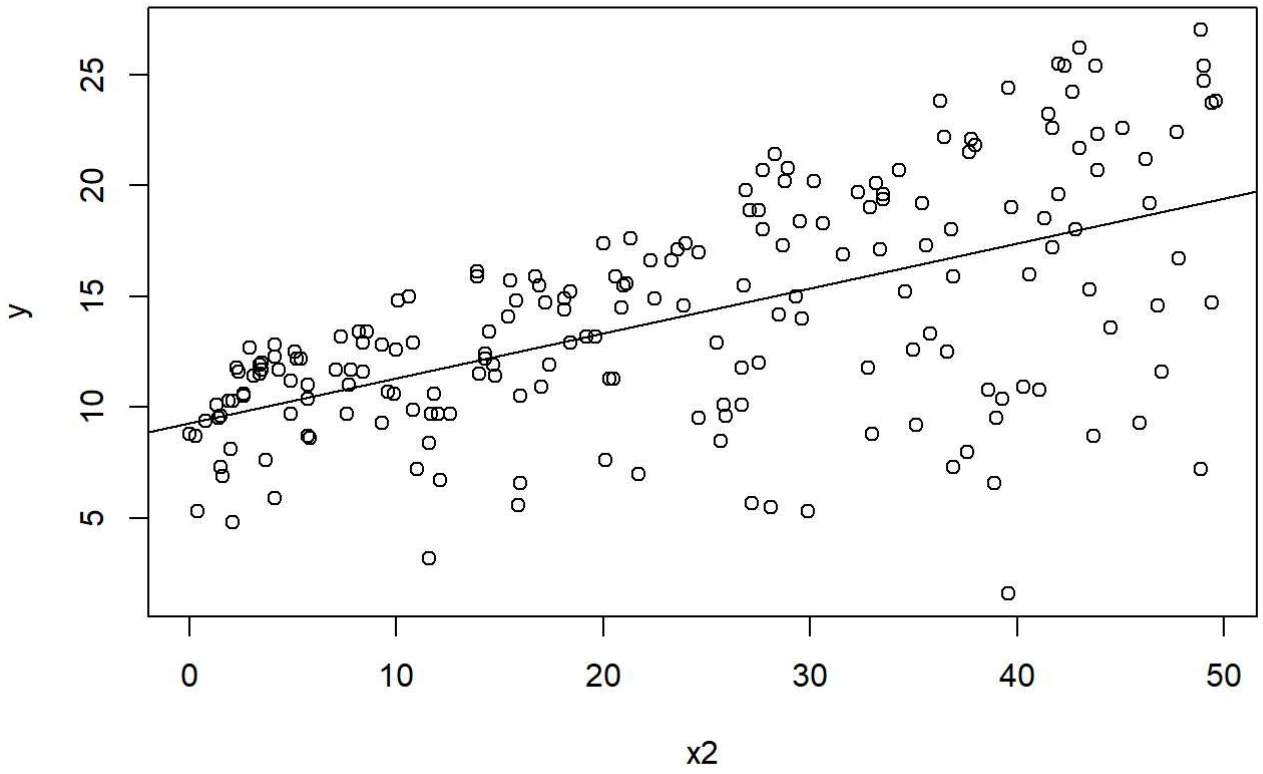



```
#using builtin regression function
#for y onto x2

model2<- lm(y ~ x2)
options(digits = 9)
summary(model2)
```

```
##
## Call:
## lm(formula = y ~ x2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -15.730471  -2.132427   0.770692   2.777527   8.181043
##
## Coefficients:
##              Estimate Std. Error  t value   Pr(>|t|)
## (Intercept) 9.3116381  0.5629005 16.54225 < 2.22e-16 ***
## x2          0.2024958  0.0204113  9.92077 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.27494 on 198 degrees of freedom
## Multiple R-squared:  0.332032,   Adjusted R-squared:  0.328659
## F-statistic: 98.4216 on 1 and 198 DF,  p-value: < 2.22e-16
```

```
plot(x2,y)
abline(model2)
```

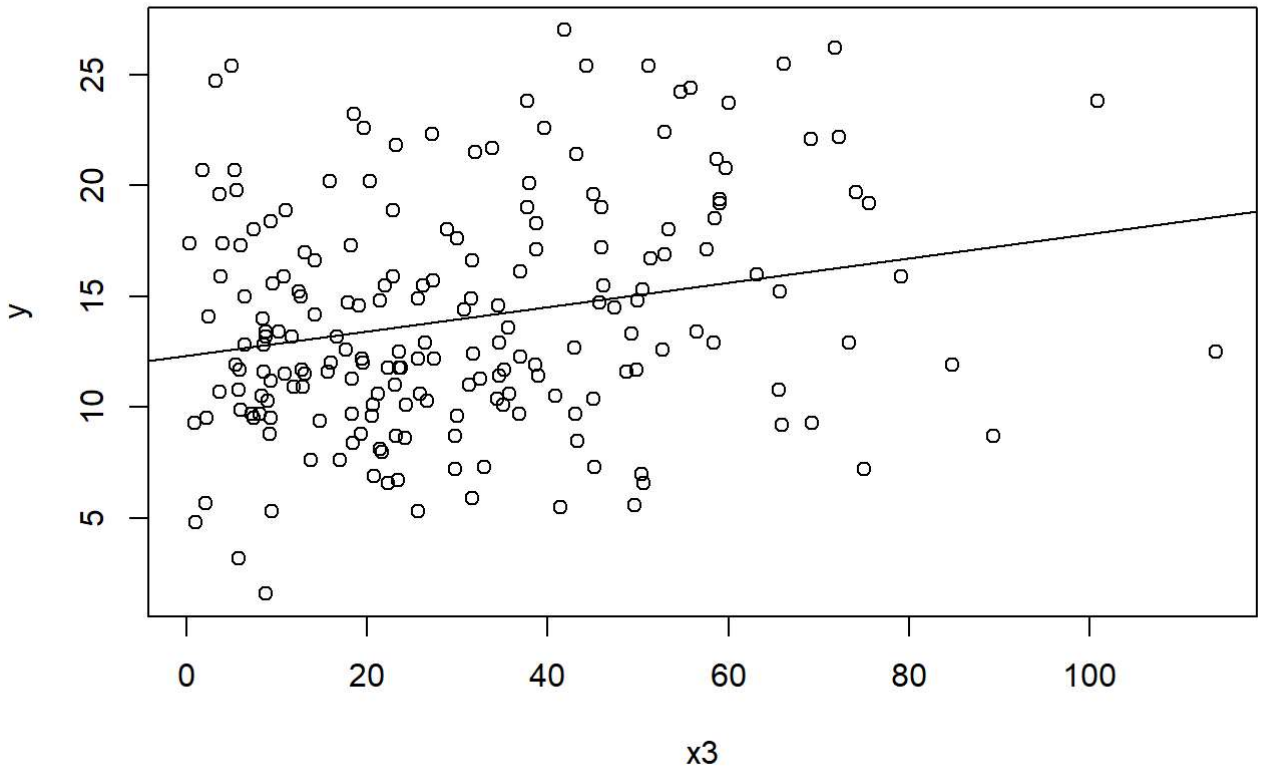



```
#using builtin regression function
#for y onto x3

model3<- lm(y ~ x3)
options(digits = 9)
summary(model3)
```

```
##
## Call:
## lm(formula = y ~ x3)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -11.227237  -3.387304  -0.839203   3.505913  12.775127
##
## Coefficients:
##               Estimate Std. Error  t value   Pr(>|t|)
## (Intercept) 12.3514071  0.6214202 19.87610 < 2.22e-16 ***
## x3           0.0546931  0.0165757  3.29959  0.0011482 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.09248 on 198 degrees of freedom
## Multiple R-squared:  0.0521204,  Adjusted R-squared:  0.0473332
## F-statistic: 10.8873 on 1 and 198 DF,  p-value: 0.0011482
```

```
plot(x3,y)
abline(model3)
```


y
5
10
15
20
25
0
20
40
60
80
100
x3