sta_project_1

Ishita Dutta, Devika Sunil Kumar, Fernanda Serna Godoy

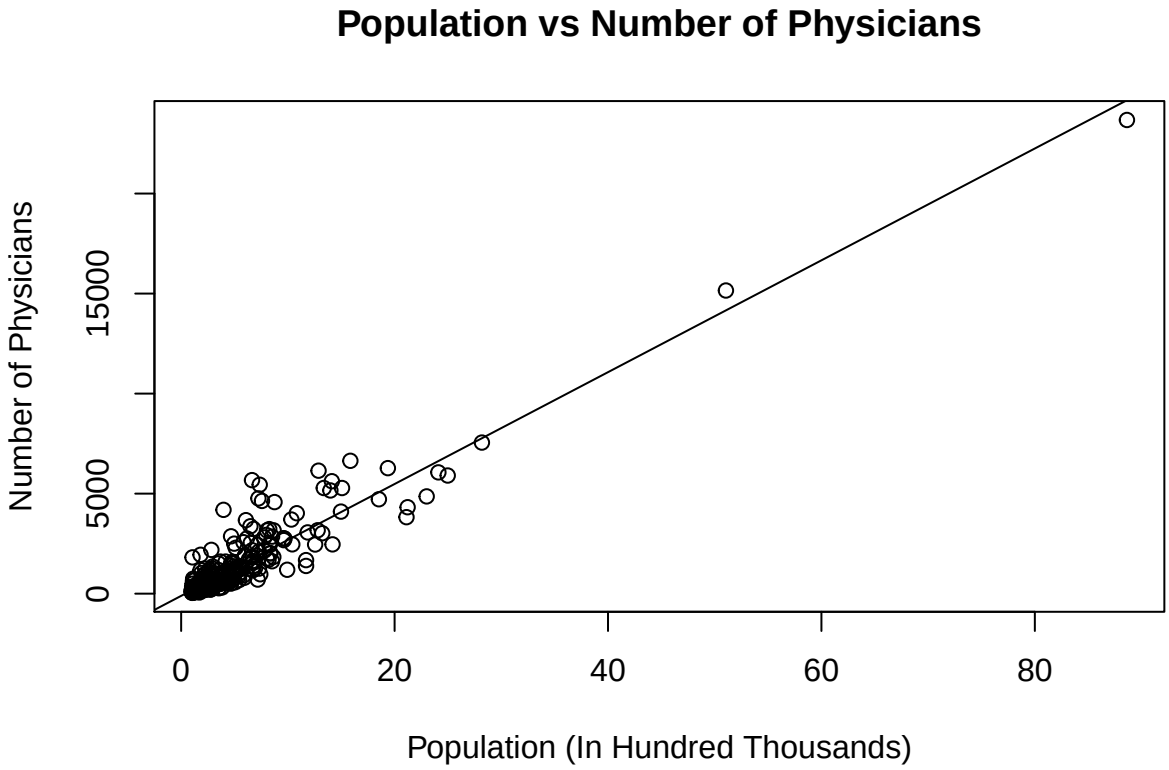
4/16/2021

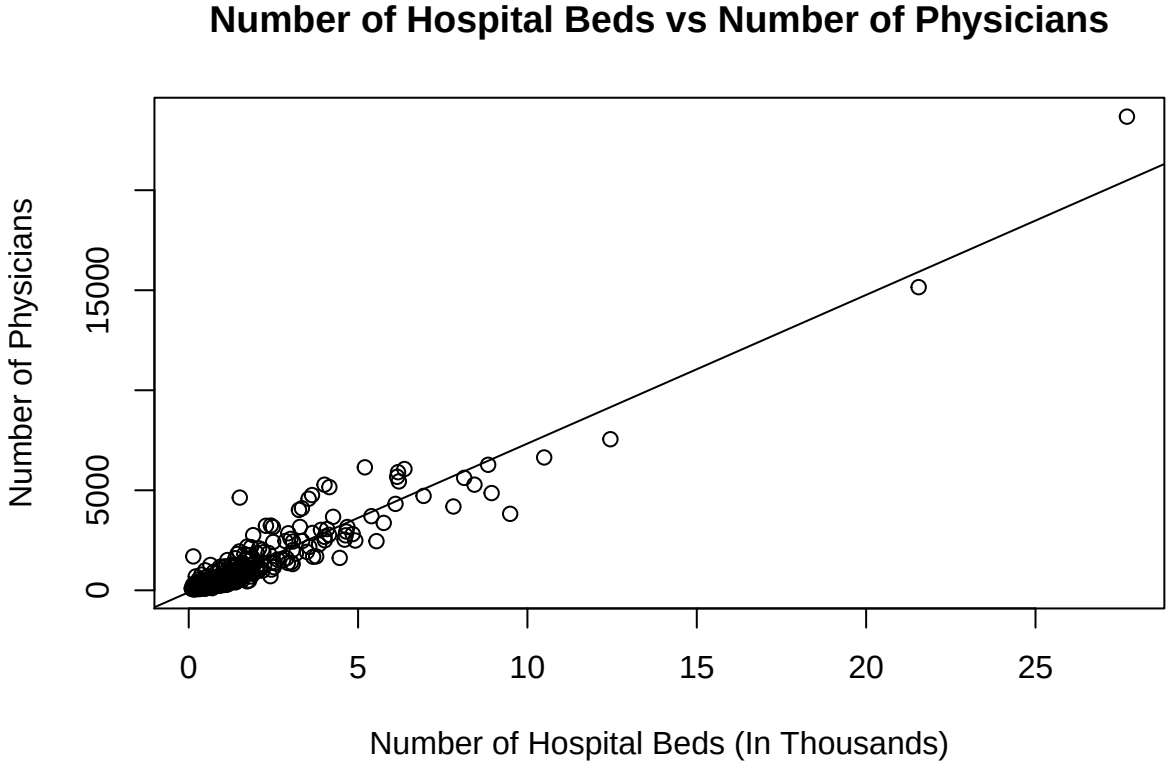
Problem 1 - Fitting Regression Models

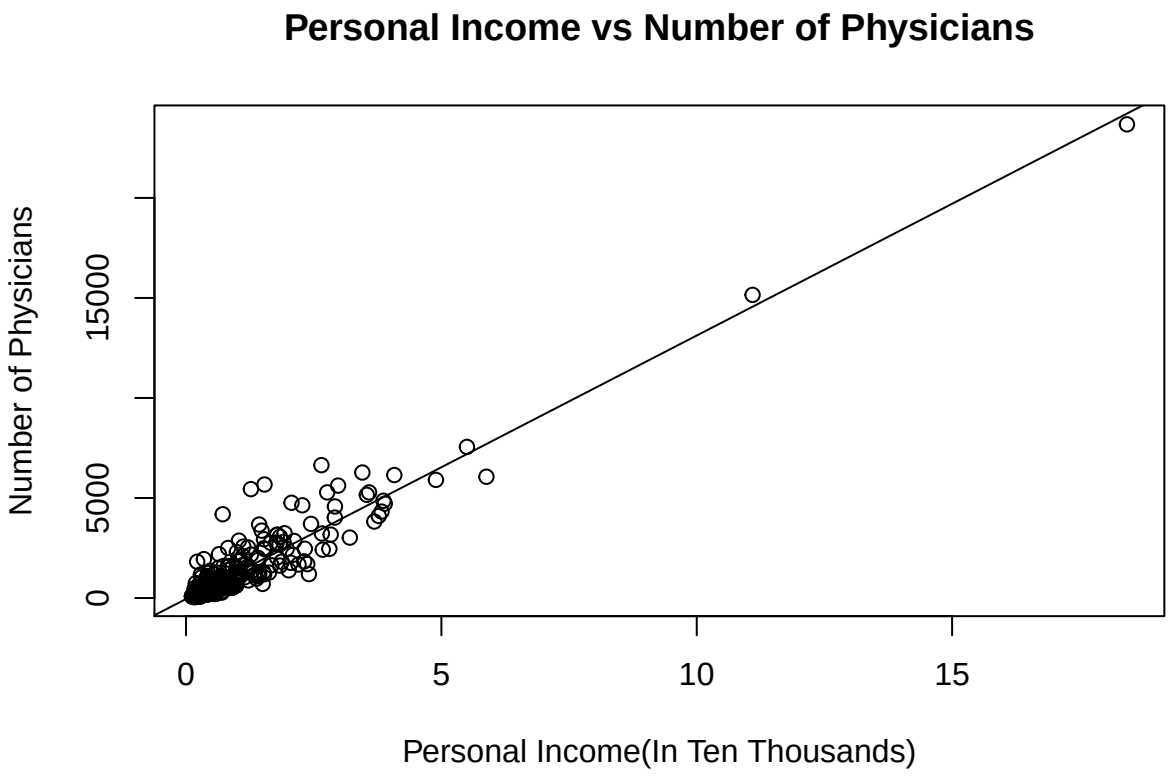
1.43 a)

```
## Population(In Hundred Thousands):

## Y =  -110.6348  +  279.5425 x

## Number of Hospital Beds(In Thousands):

## Y =  -95.93218  +  743.1164 x

## Personal Income(In Ten Thousands):

## Y =  -48.39485  +  1317.012 x
```

1.43 b)


Population vs Number of Physicians
Number of Physicians
0
5000
15000
0
20
40
60
80
Population (In Hundred Thousands)


Number of Hospital Beds vs Number of Physicians
Number of Physicians
0
5000
15000
0
5
10
15
20
25
Number of Hospital Beds (In Thousands)



1.43

calculate MSE

```
## MSE for Population Measured in Hundred Thousands:  372203.5
```

```
## MSE for Number of Hospital Beds Measured in Thousands:  11566237
```

```
## MSE for Personal Income Measured in Ten Thousands:  56681269
```

Appendix

```
knitr::opts_chunk$set(
 echo = FALSE,
 error = FALSE,
 message = FALSE,
 warning = FALSE
)
library(ggplot2)
library(viridis)
```

```
data1 = read.table("CDI.txt")
X1 = data1[,5]
X1 = X1 / 100000.0
X2 = data1[,9]
X2 = X2 / 1000.0
X3 = data1[,16]
X3 = X3 / 10000.0
Y = data1[,8]
n = length(X1)


cat("Population(In Hundred Thousands): \n")
b1hat1 = t(X1-mean(X1))%*%(Y-mean(Y))/sum((X1-mean(X1))^2)
b0hat1 = mean(Y) - b1hat1*mean(X1)
cat("Y = ", b0hat1, " + ", b1hat1, "x\n")

cat("Number of Hospital Beds(In Thousands): \n")
b1hat2 = t(X2-mean(X2))%*%(Y-mean(Y))/sum((X2-mean(X2))^2)
b0hat2 = mean(Y) - b1hat2*mean(X2)
cat("Y = ", b0hat2, " + ", b1hat2, "x\n")

cat("Personal Income(In Ten Thousands): \n")
b1hat3 = t(X3-mean(X3))%*%(Y-mean(Y))/sum((X3-mean(X3))^2)
b0hat3 = mean(Y) - b1hat3*mean(X3)
cat("Y = ", b0hat3, " + ", b1hat3, "x\n")
plot(X1, Y, xlab="Population (In Hundred Thousands)", ylab="Number of Physicians" , main = "Population
abline(lm(Y ~ X1))

plot(X2, Y, xlab="Number of Hospital Beds (In Thousands)", ylab="Number of Physicians" , main = "Number
abline(lm(Y ~ X2))

plot(X3, Y, xlab="Personal Income(In Ten Thousands)", ylab="Number of Physicians" , main = "Personal In
abline(lm(Y ~ X3))
fit.y1 = b0hat1[1] + b1hat1[1]*X1
mse1 = 1/(n-2)*sum((Y - fit.y1)^2)
cat("MSE for Population Measured in Hundred Thousands: ", mse1, "\n")

fit.y2 = b0hat2[1] + b1hat2[1]*X1
mse2 = 1/(n-2)*sum((Y - fit.y2)^2)
cat("MSE for Number of Hospital Beds Measured in Thousands: ", mse2, "\n")

fit.y3 = b0hat3[1] + b1hat3[1]*X1
mse3 = 1/(n-2)*sum((Y - fit.y3)^2)
cat("MSE for Personal Income Measured in Ten Thousands: ", mse3, "\n")
```