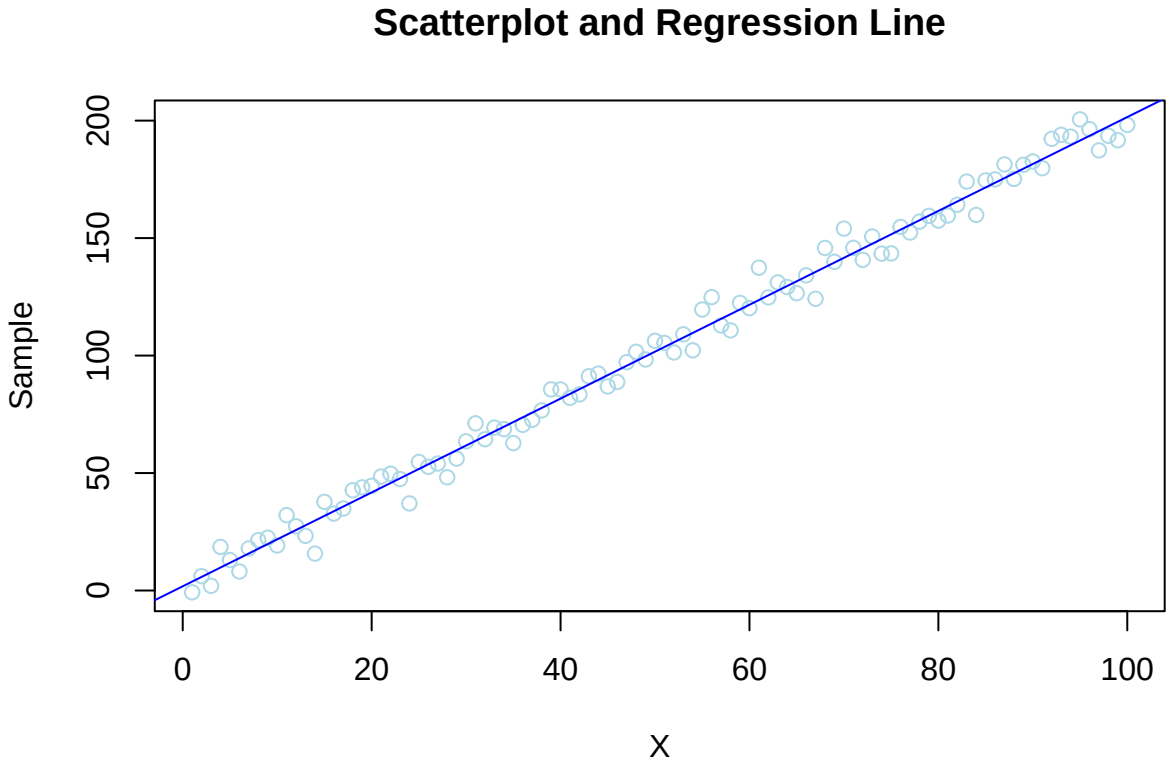
Homework8

2024-04-26

Question 1

```
set.seed(1)
x <- 1:100
epsilon <- rnorm(100,0,6)
Y_sample <- 1 + 2*x + epsilon
plot(x,Y_sample,main = "Scatterplot and Regression Line",xlab = "X", ylab = "Sample", col = "lightblue")
abline(lm(Y_sample ~ x), col = "blue")
```



Two Sided Significance Test:

Step 1: Model for data, Y~N(B0 + B1x, σ^2)

Step 2:
Null Hypothesis(H0): B1 = 2
Alternate Hypothesis(H1): B1 $\neq$ 2

Step 3: Test statistic: t statistic

Step 4:

```
B1 <- coef(lm(Y_sample ~ x))[2]
B1
```

```
##        x
## 1.997294
```

```
SE <- summary(lm(Y_sample ~ x))$coefficients[2, 2]
SE
```

```
## [1] 0.01876268
```

```
t <- (1.997294-2)/SE
t
```

```
## [1] -0.1442225
```

Step 5:
Calculating the p-value under H0,

```
p <- pt(abs(t), df = 98, lower.tail = FALSE) * 2
p
```

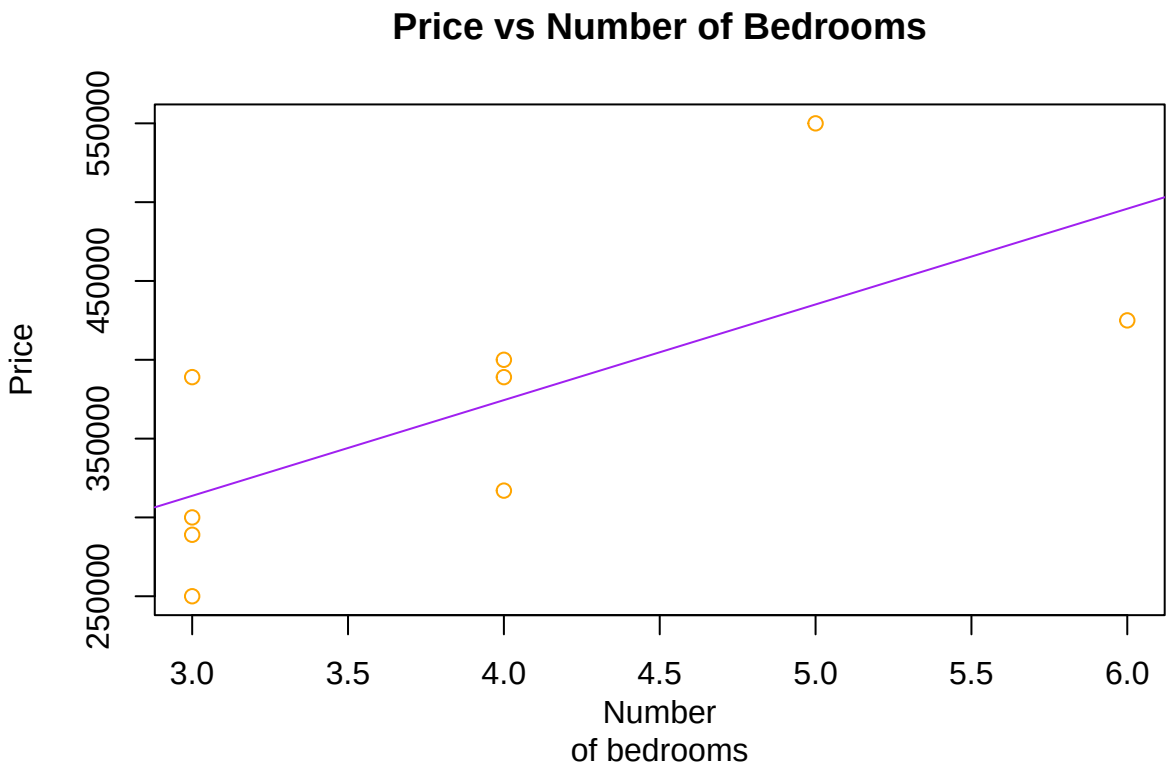
```
## [1] 0.885621
```

Step 6:
Significance level, $\alpha = 0.05$

Step 7:
p-value $= 0.8856209 > \alpha$
Hence, fail to reject the null hypothesis of B1 $= 2$

Question 2

```
price <- c(300000, 250000, 400000, 550000, 317000, 389000, 425000, 289000, 389000)
bedrooms <- c(3,3,4,5,4,3,6,3,4)
plot(bedrooms,price,main="Price vs Number of Bedrooms",xlab = "Number
of bedrooms", ylab = "Price", col = "orange")
abline(lm(price ~ bedrooms), col = "purple")
```




```
new_data <- data.frame(bedrooms = 2:8)
pred <- predict(lm(price ~ bedrooms), newdata = new_data,
        interval = c("confidence"), level = 0.95)
pred_df <- data.frame(new_data, pred)
pred_df
```

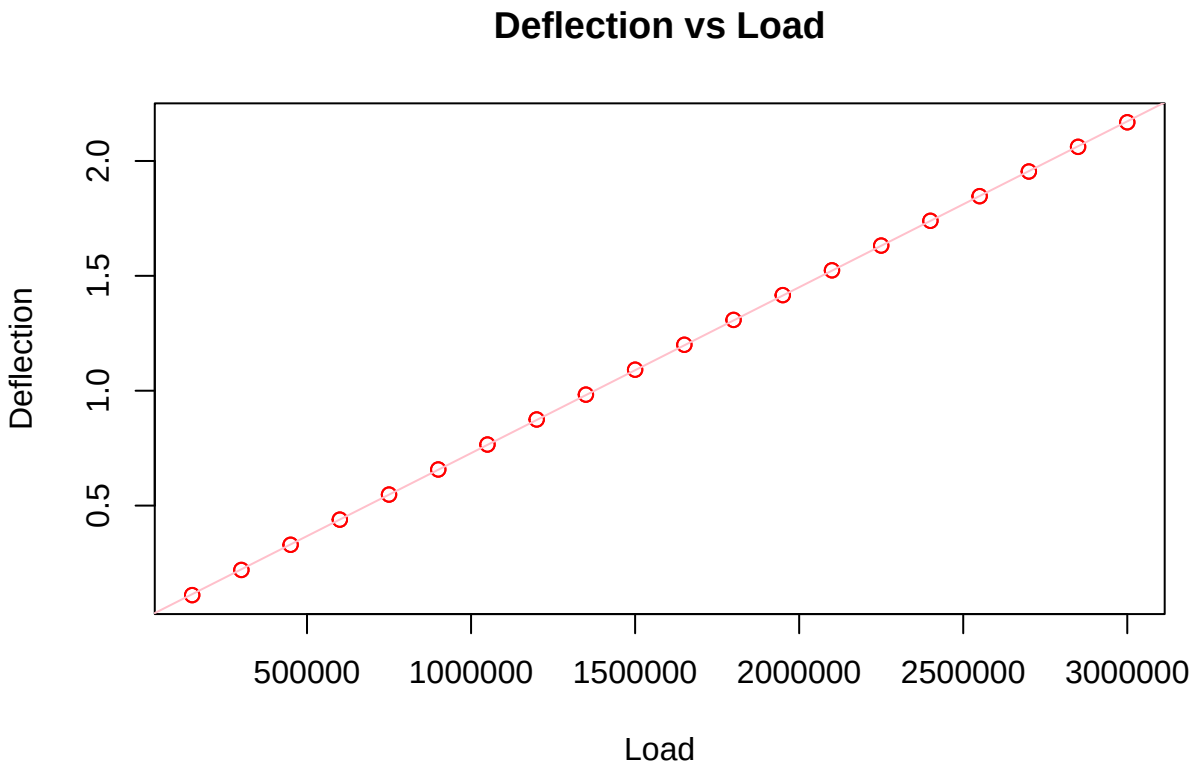
```
##   bedrooms      fit      lwr      upr
## 1        2 252987.5 136927.9 369047.1
## 2        3 313700.0 241198.2 386201.8
## 3        4 374412.5 320036.1 428788.9
## 4        5 435125.0 354065.5 516184.5
## 5        6 495837.5 368959.3 622715.7
## 6        7 556550.0 378957.5 734142.5
## 7        8 617262.5 387276.2 847248.8
```

Question 3

```
library(UsingR)
```

```
## Loading required package: MASS
## Loading required package: HistData
## Loading required package: Hmisc
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```r
data(deflection)
lm_model <- lm(Deflection ~ Load, data = deflection)
plot(deflection$Load, deflection$Deflection,
     main = "Deflection vs Load",
     xlab = "Load", ylab = "Deflection", col = "red")
abline(lm_model, col = "pink")
```




```r
confint(lm_model, level = 0.95)
```

```
##                    2.5 %       97.5 %
## (Intercept) 4.705876e-03 7.593493e-03
## Load        7.212991e-07 7.229061e-07
```