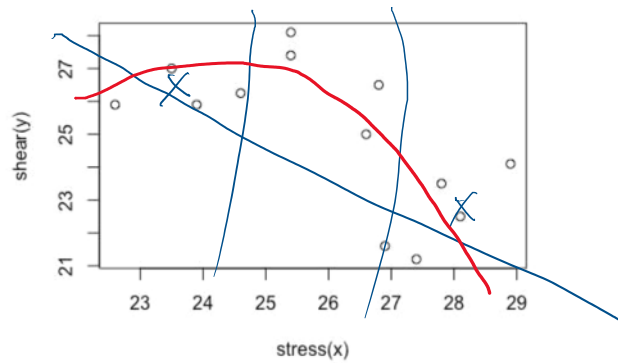


Regression Review

Question 1 - Slope

Consider the scatterplot shown below. Which statement most accurately describes the slope shown?

- ☒ 1. The slope is positive
- ☐ 2. The slope is negative
- ☐ 3. The slope is approximately zero
- ☐ 4. Cannot determine from information provided



Divide data into
 $1/3$ along x-axis

Question 2 - Y-intercept

What is the value of y-intercept (β_0)?

1. 44.022
2. -0.732
3. 7.063
4. 0.271
5. The correct answer is not provided

```
slr <- lm(y~x,data=df)
summary(slr)
```

```
##
## Call:
## lm(formula = y ~ x, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7658 -0.9534  0.1796  1.2322  2.6703
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   44.022     7.063    6.233 6.41e-05 ***
## x             -0.732     0.271   -2.701  0.0206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.813 on 11 degrees of freedom
## Multiple R-squared:  0.3987, Adjusted R-squared:  0.344
## F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064
```

$$\beta_0 = 44.0222$$

Question 3 - Slope

What is the value of the slope (β_1)?

1. 44.022

2. -0.732

3. 7.063

4. 0.271

5. The correct answer is not provided

```
s1r <- lm(y~x, data=df)
summary(s1r)
```

##

Call:

lm(formula = y ~ x, data = df)

##

Residuals:

Min 1Q Median 3Q Max

-2.7658 -0.9534 0.1796 1.2322 2.6703

##

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 44.022 7.063 6.233 6.41e-05 ***

x -0.732 0.271 -2.701 0.0206 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##

Residual standard error: 1.813 on 11 degrees of freedom

Multiple R-squared: 0.3987, Adjusted R-squared: 0.344

F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064

$$\beta_1 = -0.732$$

$$y = \beta_0 + \beta_1 x$$

Question 4 - Fitted Value

What is the predicted value of shear (y) when stress(x)=25?

1. -61.517
2. 13.838
3. 25.722
4. The correct answer is not provided

```
slr <- lm(y~x,data=df)
summary(slr)
```

```
##
## Call:
## lm(formula = y ~ x, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7658 -0.9534  0.1796  1.2322  2.6703
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   44.022      7.063   6.233 6.41e-05 ***
## x             -0.732      0.271  -2.701  0.0206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.813 on 11 degrees of freedom
## Multiple R-squared:  0.3987, Adjusted R-squared:  0.344
## F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064
```

$$\begin{aligned}\hat{y} &= \hat{\beta}_0 + \beta_1 x \\ &= 44.022 + (-0.732)(25) \\ &= 25.722\end{aligned}$$

$$e_i = y_i - \hat{y}_i = \text{actual} - \text{predicted}$$

Question 5 - Residual

What is the value of the residual when stress(x)=25.6 and strain(y)=24.9?

1. -0.1952

2. 0.1952

3. 0.3828

4. The correct answer is not provided

```
slr <- lm(y~x, data=df)
summary(slr)
```

```
##
## Call:
## lm(formula = y ~ x, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7658 -0.9534  0.1796  1.2322  2.6703
##
## Coefficients:
##      (Intercept)      44.022      7.063      6.233      6.41e-05 ***
##           x          -0.732      0.271      -2.701      0.0206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.813 on 11 degrees of freedom
## Multiple R-squared:  0.3987, Adjusted R-squared:  0.344
## F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064
```

$$= 24.9 - (\beta_0 + \beta_1 x)$$

$$= 24.9 - (44.022 + (-0.732)(25.6))$$

$$= -0.3828$$

Question 6 - Hypothesis Test on y-intercept

What is the correct decision for the hypothesis test that $H_0: \beta_0 = 0$ versus $H_1: \beta_0 \neq 0$, using $\alpha = 0.01$?

1. Fail to reject H_0
2. Reject H_0
3. Cannot determine from the information provided.

```
slr <- lm(y~x,data=df)
summary(slr)
```

```
##
## Call:
## lm(formula = y ~ x, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7658 -0.9534  0.1796  1.2322  2.6703
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   44.022      7.063    6.238 6.41e-05 ***
## x             -0.732      0.271   -2.701  0.0206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.813 on 11 degrees of freedom
## Multiple R-squared:  0.3987, Adjusted R-squared:  0.344
## F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064
```

is $p\text{-value} < \alpha = 0.01$?

is $6.4 \times 10^{-5} < 0.01$?

Yes, Reject H_0

$p\text{-value} = 6.4 \times 10^{-5}$

P-value

Question 7 - Hypothesis Test on Slope Term

What is the correct decision for the hypothesis test that $H_0: \beta_1 = 0$ versus $H_1: \beta_1 \neq 0$, using $\alpha = 0.01$?

1. Fail to reject H_0
2. Reject H_0
3. Cannot determine from the information provided.

```
slr <- lm(y~x, data=df)
summary(slr)
```

```
##
## Call:
## lm(formula = y ~ x, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7658 -0.9534  0.1796  1.2322  2.6703
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  44.022      7.063    6.233 6.41e-05 ***
## x           -0.732      0.271   -2.704  0.0206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.813 on 11 degrees of freedom
## Multiple R-squared:  0.3987, Adjusted R-squared:  0.344
## F-statistic: 7.293 on 1 and 11 DF, p-value: 0.02064
```

$$\alpha = .01$$
$$p\text{-value} = .0206$$

Is $p\text{-value} < \alpha$ or
 $.0206 < .01 \rightarrow$ No, fail to reject H_0

What if $\alpha = .05$?
is $.0206 < .05$? Yes, Reject H_0

Question 8 - Lack of Fit Test

What is the correct decision for the hypothesis test H_0 : there is no lack of fit versus H_1 : there is lack of fit, using $\alpha = 0.01$?

1. Fail to reject H_0
 2. Reject H_0
 3. Cannot determine from the information provided
- ```
library(EnvStats)
```

```
##
Attaching package: 'EnvStats'
The following objects are masked from 'package:stats':
predict, predict.lm
anovaPE(slr)
```

|             | Df | Sum Sq | Mean Sq | F value | Pr(>F)  |
|-------------|----|--------|---------|---------|---------|
| x           | 1  | 23.965 | 23.9649 | 97.816  | 0.06415 |
| Lack of Fit | 10 | 35.902 | 3.5902  | 14.654  | 0.20078 |
| Pure Error  | 1  | 0.245  | 0.2450  |         |         |

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

last column  
is p-value

$$p\text{-value} = 0.20078$$
$$\alpha = 0.01$$

Is ~~the~~  $p\text{-value} < \alpha$  or  
 $0.2 < 0.01$ ? No, therefore  
we fail to reject  $H_0$

$$\alpha = 0.01$$

$H_0$ : Good Model  
 $H_1$ : Bad Model