Basic Statistics Review

Quiz, 9 questions

1 point

1.

Enter the following dataset in R using concatenation operator. You may edit the code fragment below:

37, 86, 79, 95, 61, 93, 19, 98, 121, 26, 39, 11, 26, 75, 29,130, 42, 8###

Obtain 5-number summary. You may edit the code fragment below. What is the sample mean?

```
1  data=c(37, 86, 79, 95, 61, 93, 19, 98, 121, 26, 39, 11, 26, 75, 29,130,
     42, 8) # Edit this line
2  summary(data) # Edit this line
```

Run

Reset

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   8.00   26.75   51.50   59.72   91.25  130.00
```

- (●) 59.72
- ( ) 51.50

1 point

2.

Basic Statistics Review

Quiz, 9 questions

Find the summary of the dataset given in the following code block. What is the 3rd quartile?

```
1  data=c(37, 86, 79, 95, 61, 93, 19, 98, 121, 26, 39, 11, 26, 75, 29,130,
     42, 8) # Edit this line
2  summary(data) # Edit this line
```

Run

Reset

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   8.00   26.75   51.50   59.72   91.25  130.00
```

- 130
- 91.25
- 26.75

1 point

3.

Basic Statistics Review

Quiz, 9 questions

We look at the dataset titled 'cheddar' from 'faraway' package. Continue the code in the following code block to look at the description of the dataset 'cheddar' using help() routine. How many observations and variables are there in the dataset?

```
1  library(faraway)
2  help(cheddar)
3
```

Run

Reset

```
cheddar                package:faraway                R Documentation

_✓T_✓a_✓s_✓t_✓e _✓o_✓f _✓C_✓h_✓e_✓d_✓d_✓a_✓r _✓c_✓h_✓e_✓e_✓s_✓e

_✓D_✓e_✓s_✓c_✓r_✓i_✓p_✓t_✓i_✓o_✓n:

     In a study of cheddar cheese from the LaTrobe Valley of Victoria,
     Australia, samples of cheese were analyzed for their chemical
     composition and were subjected to taste tests. Overall taste
     scores were obtained by combining the scores from several tasters.

_✓U_✓s_✓a_✓g_✓e:

     data(cheddar)

_✓F_✓o_✓r_✓m_✓a_✓t:

     A data frame with 30 observations on the following 4 variables.

     'taste' a subjective taste score

     'Acetic' concentration of acetic acid (log scale)

     'H2S' concentration of hydrogen sulfice (log scale)

     'Lactic' concentration of lactic acid

_✓S_✓o_✓u_✓r_✓c_✓e:

     Unknown

_✓E_✓x_✓a_✓m_✓p_✓l_✓e_✓s:

     data(cheddar)
     ## maybe str(cheddar) ; plot(cheddar) ...
```

- 4 observations and 30 variables
- 30 observations and 5 variables
- 30 Observations and 4 variables

Basic Statistics Review

Quiz, 9 questions

1 point

4.

We are still working on the dataset 'cheddar' from the package 'faraway'. Apply simple linear regression model for the bivariate data 'taste' (modeled as a random variable Y) vs 'H2S' (modeled as a random variable X) in the dataset 'cheddar' in using lm() routine in the following code block. What is the model?

```
1  library(faraway)
2  m=lm(taste~H2S, data=cheddar)
3  summary(m)
4
```

Run

Reset

```
Call:
lm(formula = taste ~ H2S, data = cheddar)

Residuals:
    Min      1Q  Median      3Q     Max
-15.426  -7.611  -3.491   6.420  25.687

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -9.7868     5.9579  -1.643    0.112
H2S           5.7761     0.9458   6.107 1.37e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.83 on 28 degrees of freedom
Multiple R-squared:  0.5712,    Adjusted R-squared:  0.5558
F-statistic: 37.29 on 1 and 28 DF,  p-value: 1.374e-06
```

- $Y = -1.643 + 6.107 * X + \epsilon$

  where $\epsilon \sim N(0, 10.83^2)$.

- $Y = -9.7868 + 5.7761 * X + \epsilon$

  where $\epsilon \sim N(0, 10.83^2)$.

$Y = --9.7868 + 5.7761 * X + \epsilon$

Basic Statistics Review

where $\epsilon \sim N(0, 0.5712^2)$.

Quiz, 9 questions

1 point

5.

What is the sum of the residuals in the simple linear regression model of Question 4?

```
1  library(faraway)
2  m=lm(taste~H2S, data=cheddar)
3  res.m = resid(m)
4  sum(res.m)
```

Run

Reset

```
[1] 1.332268e-15
```

- (selected) 0
- 736

1 point

6.

What is the sum of the fitted values in the simple linear regression model of Question 4? We can get the fitted values by using lm()$fitted routine.

```
1  library(faraway)
2  m=lm(taste~H2S, data=cheddar)
3  sum(m$fitted)
```

Run

Reset

```
[1] 736
```

- 0
- (selected) 736

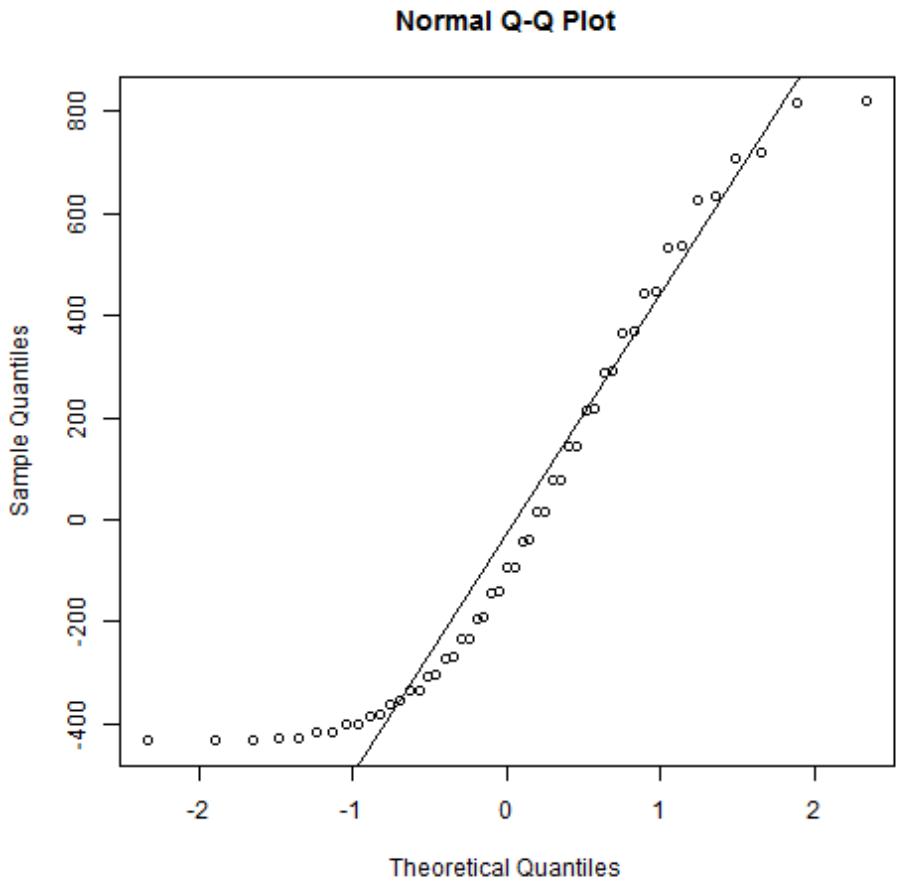
Basic Statistics Review

Quiz, 9 questions

1 point

7.

Does this data set appear to be normally distributed?



- Yes.
- No.

1 point

8.

Basic Statistics Review

Quiz, 9 questions

Suppose you are testing the null hypothesis that a population mean is 0 against the alternative that it is not zero at the alpha=0.05 level of significance.

Given the following function call and printout, can you reject your null hypothesis?

```
1  data = c(7,  5,  1,  7,  2,  5,  2,  4, 10,  6);
2  t.test(data, alternative = "two.sided", paired=FALSE)
```

Run

Reset

```
        One Sample t-test

data:  data
t = 5.6003, df = 9, p-value = 0.0003342
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 2.920702 6.879298
sample estimates:
mean of x
      4.9
```

- ○ No.
- ◉ Yes.

1 point

9.

Do you believe the R printout matches the regression in the figure?

Basic Statistics Review

Quiz, 9 questions

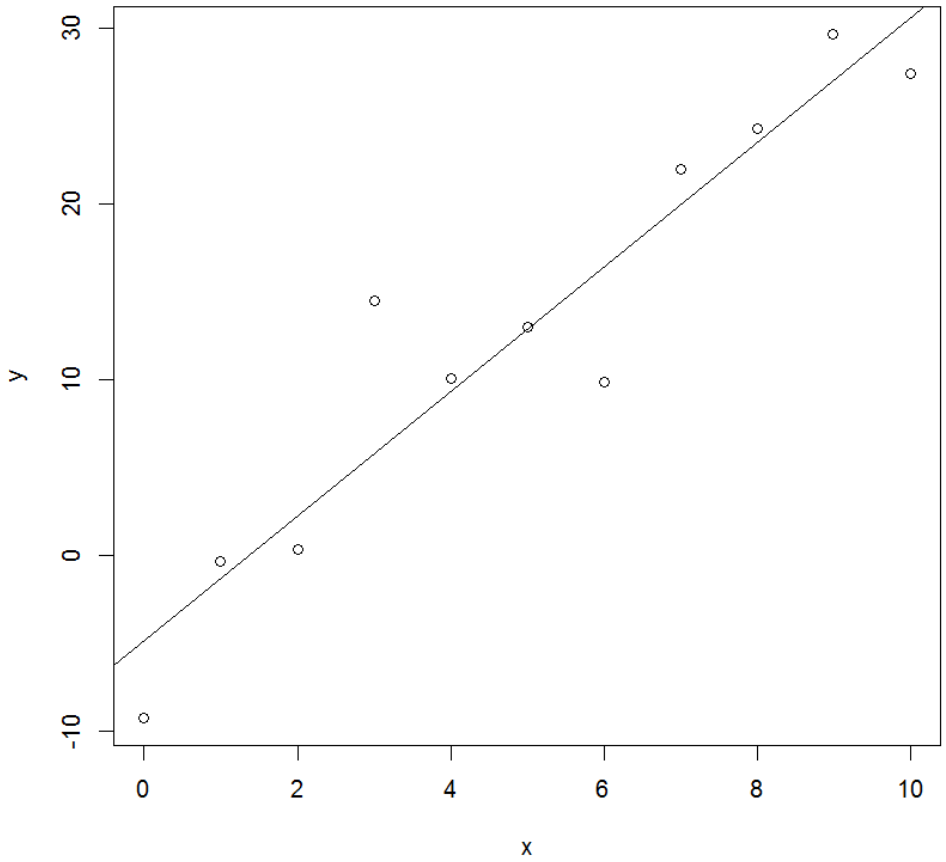
Call:

lm(formula = y ~ x)

Coefficients:

(Intercept) x

-4.48 -2.82



- Yes.
- No

I, **Mark R. Lytell**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Basic Statistics Review

Quiz, 9 questions

Submit Quiz