Review of Regression
Sections 26.1, 26.2

Lecture 50

Robb T. Koether

Hampden-Sydney College

Wed, Apr 20, 2016

Outline

Outline

Review of Regression

- Way back in Chapter 5 we learned how to calculate the equation of the regression line
$$\hat{y} = a + bx.$$
- x is the explanatory variable and y is the response variable.
- We also learned how to calculate the correlation r that measures the strength and direction of the relation.
- But we did not address an important question:

Review of Regression

- Way back in Chapter 5 we learned how to calculate the equation of the regression line
$$\hat{y} = a + bx.$$
- x is the explanatory variable and y is the response variable.
- We also learned how to calculate the correlation r that measures the strength and direction of the relation.
- But we did not address an important question:

Is there *really* any relation at all?

Example

Example (Quiz Average vs. Test Average)

- Is there a linear relation between quiz averages and test averages?

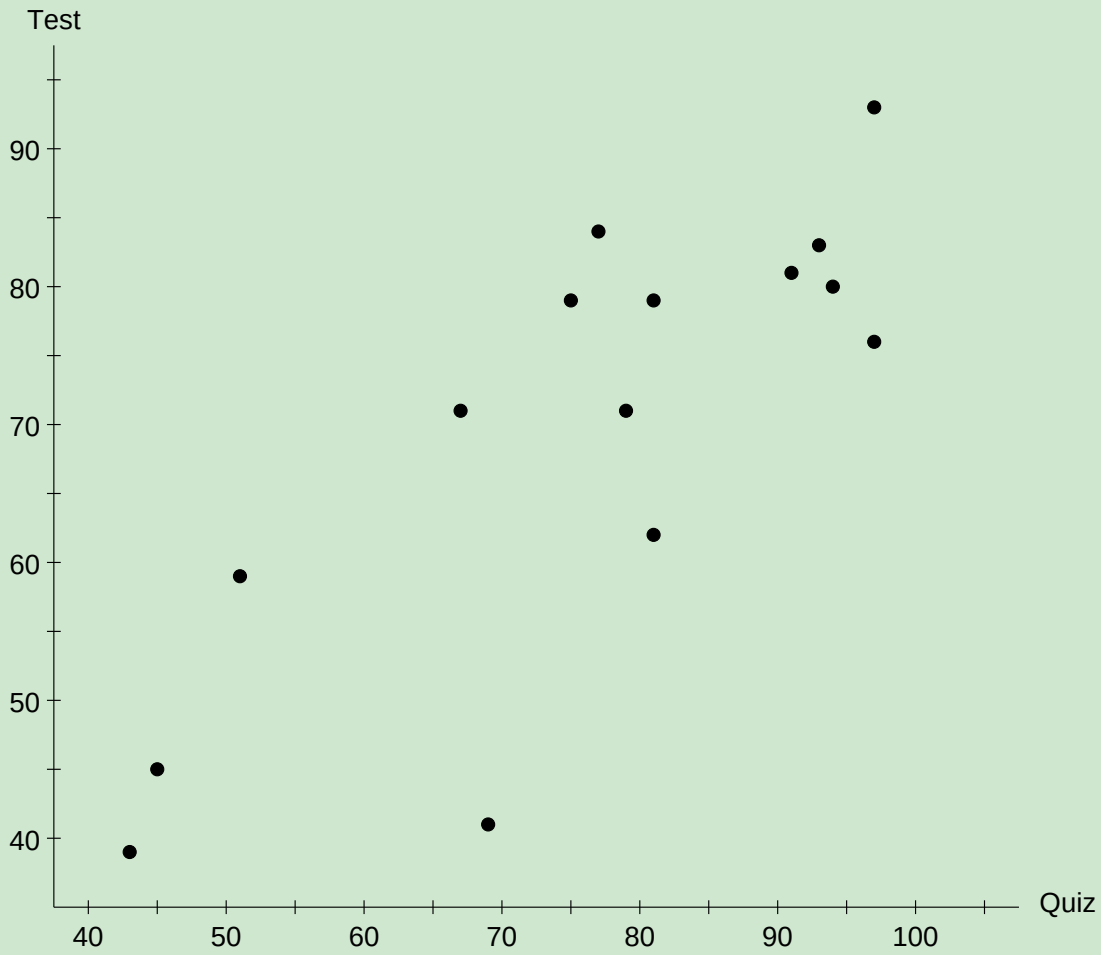
Quiz	Test
79	71
45	45
51	59
81	62
69	41

Quiz	Test
75	79
97	93
97	76
43	39
67	71

Quiz	Test
94	80
77	84
81	79
91	81
93	93

Example

Example (Quiz Average vs. Test Average)



Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.

Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.
- Then calculate separately

Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.
- Then calculate separately

$$\sum (x_i - \overline{x})^2,$$

Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.
- Then calculate separately

$$\sum (x_i - \overline{x})^2,$$
$$\sum (y_i - \overline{y})^2,$$

Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.
- Then calculate separately

$$\sum (x_i - \overline{x})^2,$$
$$\sum (y_i - \overline{y})^2,$$
$$\sum (x_i - \overline{x})(y_i - \overline{y}).$$

Correlation

- To find the correlation r, first, calculate $\sum x_i$ and $\sum y_i$ to get $\overline{x}$ and $\overline{y}$.
- Then calculate separately

$$\sum (x_i - \overline{x})^2,$$
$$\sum (y_i - \overline{y})^2,$$
$$\sum (x_i - \overline{x})(y_i - \overline{y}).$$

- Then calculate

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}.$$

The Regression Line

- Recall from Chapter 5 that the regression equation is of the form

$$\hat{y} = a + bx.$$

- Let s_x and s_y be the standard deviations of x and y, respectively.

$$s_x = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}},$$

$$s_y = \sqrt{\frac{\sum (y_i - \overline{y})^2}{n-1}},$$

The Regression Line

- Then

$$b = r\left(\frac{s_y}{s_x}\right),$$
$$a = \overline{y} - b\overline{x}.$$

Example

Example (Quiz Average vs. Test Average)

- In the example of the quiz and test data, we find

$$\sum x_i = 1139,$$
$$\sum y_i = 1054,$$
$$n = 15.$$

Example

Example (Quiz Average vs. Test Average)

- In the example of the quiz and test data, we find

$$\sum x_i = 1139,$$
$$\sum y_i = 1054,$$
$$n = 15.$$

- We calculate

$$\overline{x} = \frac{1140}{15} = 76.0,$$
$$\overline{y} = \frac{1053}{15} = 70.2.$$

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71					
45	45					
51	59					
81	62					
69	41					
75	79					
97	93					
97	76					
43	39					
67	71					
94	80					
77	84					
81	79					
91	81					
93	93					

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71					
45	45					
51	59					
81	62					
69	41					
75	79					
97	93					
97	76					
43	39					
67	71					
94	80					
77	84					
81	79					
91	81					
93	93					
1140						

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71					
45	45					
51	59					
81	62					
69	41					
75	79					
97	93					
97	76					
43	39					
67	71					
94	80					
77	84					
81	79					
91	81					
93	93					
1140	1053					

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0				
45	45	-31.0				
51	59	-25.0				
81	62	5.0				
69	41	-7.0				
75	79	-1.0				
97	93	21.0				
97	76	21.0				
43	39	-33.0				
67	71	-9.0				
94	80	18.0				
77	84	1.0				
81	79	5.0				
91	81	15.0				
93	93	17.0				
1140	1053					

Example

Example (Quiz Average vs. Test Average)

x	y	$x - \overline{x}$	$y - \overline{y}$	$(x - \overline{x})^2$	$(y - \overline{y})^2$	$(x - \overline{x})(y - \overline{y})^2$
79	71	3.0				
45	45	-31.0				
51	59	-25.0				
81	62	5.0				
69	41	-7.0				
75	79	-1.0				
97	93	21.0				
97	76	21.0				
43	39	-33.0				
67	71	-9.0				
94	80	18.0				
77	84	1.0				
81	79	5.0				
91	81	15.0				
93	93	17.0				
1140	1053	0.0				

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8			
45	45	-31.0	-25.2			
51	59	-25.0	-11.2			
81	62	5.0	-8.2			
69	41	-7.0	-29.2			
75	79	-1.0	8.8			
97	93	21.0	22.8			
97	76	21.0	5.8			
43	39	-33.0	-31.2			
67	71	-9.0	0.8			
94	80	18.0	9.8			
77	84	1.0	13.8			
81	79	5.0	8.8			
91	81	15.0	10.8			
93	93	17.0	22.8			
1140	1053	0.0				

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8			
45	45	-31.0	-25.2			
51	59	-25.0	-11.2			
81	62	5.0	-8.2			
69	41	-7.0	-29.2			
75	79	-1.0	8.8			
97	93	21.0	22.8			
97	76	21.0	5.8			
43	39	-33.0	-31.2			
67	71	-9.0	0.8			
94	80	18.0	9.8			
77	84	1.0	13.8			
81	79	5.0	8.8			
91	81	15.0	10.8			
93	93	17.0	22.8			
1140	1053	0.0	0.0			

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00		
45	45	-31.0	-25.2	961.00		
51	59	-25.0	-11.2	625.00		
81	62	5.0	-8.2	25.00		
69	41	-7.0	-29.2	49.00		
75	79	-1.0	8.8	1.00		
97	93	21.0	22.8	441.00		
97	76	21.0	5.8	441.00		
43	39	-33.0	-31.2	1089.00		
67	71	-9.0	0.8	81.00		
94	80	18.0	9.8	324.00		
77	84	1.0	13.8	1.00		
81	79	5.0	8.8	25.00		
91	81	15.0	10.8	225.00		
93	93	17.0	22.8	289.00		
1140	1053	0.0	0.0			

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00		
45	45	-31.0	-25.2	961.00		
51	59	-25.0	-11.2	625.00		
81	62	5.0	-8.2	25.00		
69	41	-7.0	-29.2	49.00		
75	79	-1.0	8.8	1.00		
97	93	21.0	22.8	441.00		
97	76	21.0	5.8	441.00		
43	39	-33.0	-31.2	1089.00		
67	71	-9.0	0.8	81.00		
94	80	18.0	9.8	324.00		
77	84	1.0	13.8	1.00		
81	79	5.0	8.8	25.00		
91	81	15.0	10.8	225.00		
93	93	17.0	22.8	289.00		
1140	1053	0.0	0.0	4586.00		

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00	0.64	
45	45	-31.0	-25.2	961.00	635.04	
51	59	-25.0	-11.2	625.00	125.44	
81	62	5.0	-8.2	25.00	67.24	
69	41	-7.0	-29.2	49.00	852.64	
75	79	-1.0	8.8	1.00	77.44	
97	93	21.0	22.8	441.00	519.84	
97	76	21.0	5.8	441.00	33.64	
43	39	-33.0	-31.2	1089.00	973.44	
67	71	-9.0	0.8	81.00	0.64	
94	80	18.0	9.8	324.00	96.04	
77	84	1.0	13.8	1.00	190.44	
81	79	5.0	8.8	25.00	77.44	
91	81	15.0	10.8	225.00	116.64	
93	93	17.0	22.8	289.00	519.84	
1140	1053	0.0	0.0	4586.00		

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00	0.64	
45	45	-31.0	-25.2	961.00	635.04	
51	59	-25.0	-11.2	625.00	125.44	
81	62	5.0	-8.2	25.00	67.24	
69	41	-7.0	-29.2	49.00	852.64	
75	79	-1.0	8.8	1.00	77.44	
97	93	21.0	22.8	441.00	519.84	
97	76	21.0	5.8	441.00	33.64	
43	39	-33.0	-31.2	1089.00	973.44	
67	71	-9.0	0.8	81.00	0.64	
94	80	18.0	9.8	324.00	96.04	
77	84	1.0	13.8	1.00	190.44	
81	79	5.0	8.8	25.00	77.44	
91	81	15.0	10.8	225.00	116.64	
93	93	17.0	22.8	289.00	519.84	
1140	1053	0.0	0.0	4586.00	4286.40	

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00	0.64	2.4
45	45	-31.0	-25.2	961.00	635.04	781.2
51	59	-25.0	-11.2	625.00	125.44	280.0
81	62	5.0	-8.2	25.00	67.24	-41.0
69	41	-7.0	-29.2	49.00	852.64	204.4
75	79	-1.0	8.8	1.00	77.44	-8.8
97	93	21.0	22.8	441.00	519.84	478.8
97	76	21.0	5.8	441.00	33.64	121.8
43	39	-33.0	-31.2	1089.00	973.44	1029.6
67	71	-9.0	0.8	81.00	0.64	-7.2
94	80	18.0	9.8	324.00	96.04	176.4
77	84	1.0	13.8	1.00	190.44	13.8
81	79	5.0	8.8	25.00	77.44	44.0
91	81	15.0	10.8	225.00	116.64	162.0
93	93	17.0	22.8	289.00	519.84	387.6
1140	1053	0.0	0.0	4586.00	4286.40	

Example

Example (Quiz Average vs. Test Average)

x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})^2$	$(y-\overline{y})^2$	$(x-\overline{x})(y-\overline{y})^2$
79	71	3.0	0.8	9.00	0.64	2.4
45	45	-31.0	-25.2	961.00	635.04	781.2
51	59	-25.0	-11.2	625.00	125.44	280.0
81	62	5.0	-8.2	25.00	67.24	-41.0
69	41	-7.0	-29.2	49.00	852.64	204.4
75	79	-1.0	8.8	1.00	77.44	-8.8
97	93	21.0	22.8	441.00	519.84	478.8
97	76	21.0	5.8	441.00	33.64	121.8
43	39	-33.0	-31.2	1089.00	973.44	1029.6
67	71	-9.0	0.8	81.00	0.64	-7.2
94	80	18.0	9.8	324.00	96.04	176.4
77	84	1.0	13.8	1.00	190.44	13.8
81	79	5.0	8.8	25.00	77.44	44.0
91	81	15.0	10.8	225.00	116.64	162.0
93	93	17.0	22.8	289.00	519.84	387.6
1140	1053	0.0	0.0	4586.00	4286.40	3625.0

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$s_x = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$s_x = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}}$$
$$= \sqrt{\frac{4586}{14}}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$\begin{aligned} s_x &= \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}} \\ &= \sqrt{\frac{4586}{14}} \\ &= 18.10. \end{aligned}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$s_x = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}}$$
$$= \sqrt{\frac{4586}{14}}$$
$$= 18.10.$$
$$s_y = \sqrt{\frac{\sum (y_i - \overline{y})^2}{n-1}}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$
\begin{aligned}
s_x &= \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}} \\
&= \sqrt{\frac{4586}{14}} \\
&= 18.10.
\end{aligned}
$$

$$
\begin{aligned}
s_y &= \sqrt{\frac{\sum (y_i - \overline{y})^2}{n-1}} \\
&= \sqrt{\frac{4286.4}{14}}
\end{aligned}
$$

Example

Example (Quiz Average vs. Test Average)

- Calculate s_x and s_y:

$$
\begin{aligned}
s_x &= \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}} \\
&= \sqrt{\frac{4586}{14}} \\
&= 18.10.
\end{aligned}
$$

$$
\begin{aligned}
s_y &= \sqrt{\frac{\sum (y_i - \overline{y})^2}{n-1}} \\
&= \sqrt{\frac{4286.4}{14}} \\
&= 17.50.
\end{aligned}
$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$
$$= \frac{3625}{\sqrt{(4586)(4286.4)}}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right)$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right) = (0.8176)\left(\frac{17.50}{18.10}\right)$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right) = (0.8176)\left(\frac{17.50}{18.10}\right) = 0.7905$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right) = (0.8176)\left(\frac{17.50}{18.10}\right) = 0.7905$$

$$a = \overline{y} - b\overline{x}$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right) = (0.8176)\left(\frac{17.50}{18.10}\right) = 0.7905$$

$$a = \overline{y} - b\overline{x} = 70.2 - (0.7905)(76.0)$$

Example

Example (Quiz Average vs. Test Average)

- Calculate r:

$$\begin{aligned} r &= \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}} \\ &= \frac{3625}{\sqrt{(4586)(4286.4)}} \\ &= 0.8176. \end{aligned}$$

- Then calculate a and b:

$$b = r\left(\frac{s_y}{s_x}\right) = (0.8176)\left(\frac{17.50}{18.10}\right) = 0.7905$$

$$a = \overline{y} - b\overline{x} = 70.2 - (0.7905)(76.0) = 10.12.$$

Example

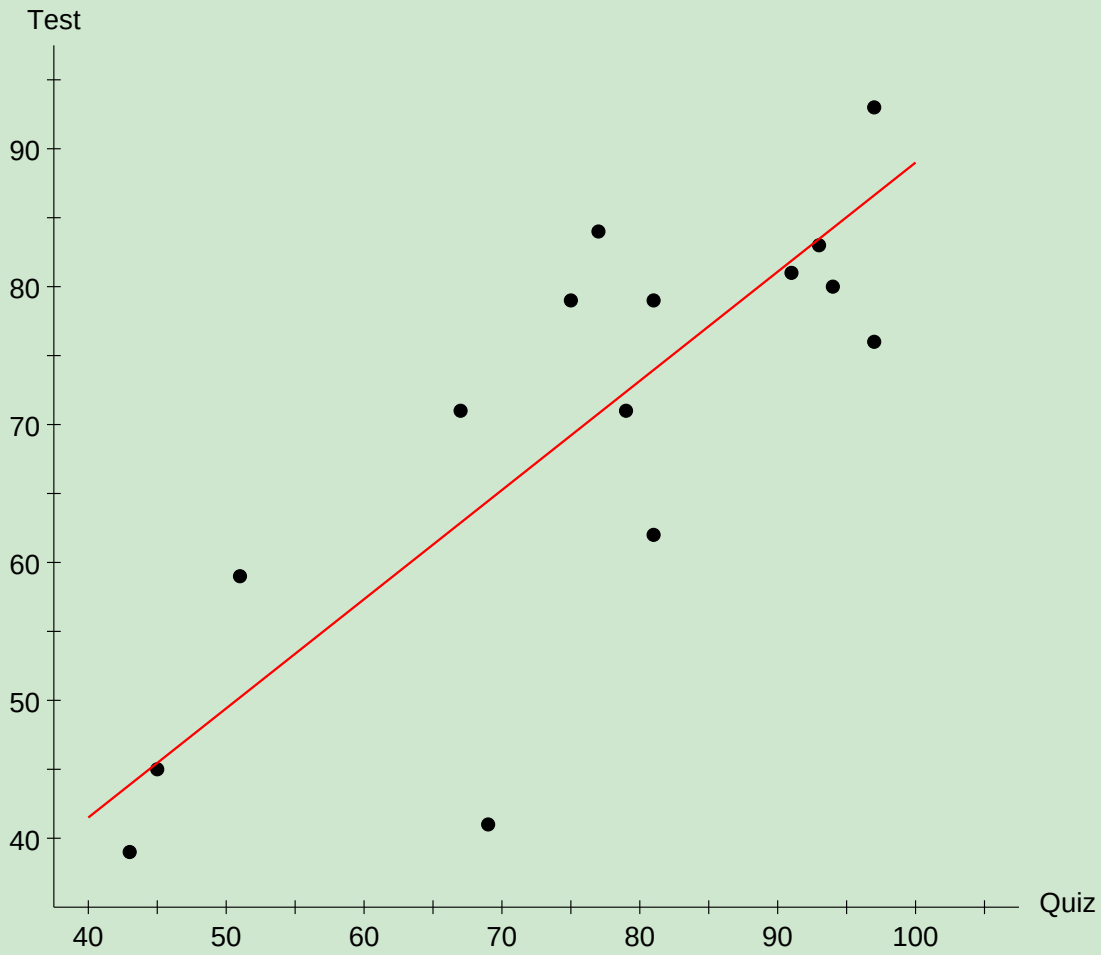
Example (Quiz Average vs. Test Average)

- Finally, we have the regression equation:

$$\hat{y} = 10.12 + 0.7905x.$$

Example

Example (Quiz Average vs. Test Average)



Outline

Assignment

Assignment

- Read Sections 26.1, 26.2.
- Apply Your Knowledge: 1.
- Check Your Skills: 16, 17, 18.
- Exercises 25(a)(b).