

## Notes 9.1 & 9.2 Correlation & Linear Regression

### Correlation

- A correlation is a comparison between 2 quantities
- The data can be represented by  $(x, y)$  where  $x$  is the independent input variable, and  $y$  is the dependent or output variable.
  - Domain (under independent input)
  - Range (under output)
- One way to determine whether a linear correlation exists between two variables is to use a \_\_\_\_\_.

• Correlation Coefficient : Linear Regression "Test"  
 how well do the data pts. make a line

Example 1: "r"

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following

Monday are given below:

0 to 1 or 0 to -1  
 positive ↗ or negative ↘

Hours spent online, $x$	0	1	2	3	3	5	5	5	6	7	7	10
Test score, $y$	96	85	82	74	95	68	76	84	58	65	75	50

- a. Sketch a graph of the distribution and then describe the distribution.



- b. Find the regression line.

- c. Find the correlation coefficient.

- d. Use the regression line to predict the test scores given the time online:

$x = 4$  hours

$x = 9$  hours

$x = 15$  hours

## Notes 9.1 & 9.2 Correlation & Linear Regression

### Example 2:

The budgets & worldwide grosses of 15 of the most expensive 20<sup>th</sup> Century Fox Movies are shown.

<b>Budget, x</b> (millions)	200	150	125	125	115	115	115	110	110	110	105	102	100	100	100
<b>Gross, y</b> (millions)	1835.4	459.4	406.4	542.7	924.3	656.7	848.5	571.1	211.4	150.5	348.8	358.8	365.3	359.1	249.0

- a. Sketch a graph of the distribution and then describe the distribution.



- b. Find the regression line.
- c. Find the correlation coefficient.
- d. Use the regression line to predict the gross amount of money for the given budget:

*\$120,000,000*

*\$93,000,000*

2nd + Catalog then... scroll down to DiagnosticON

Press Enter to get the words "Done"

Input x into L1      Input y into L2

To get Equation & "r" Stat + scroll over to Calc #4 LinReg (ax + b)

## Notes 9.1 & 9.2 Correlation & Linear Regression

### Correlation

- A correlation is a comparison of 2 quantities.
- The data can be represented by  $(x, y)$  where  $x$  is the independent input variable, and  $y$  is the dependent output variable.   
domain range
- One way to determine whether a linear correlation exists between two variables is to use a Linear Regression

"the line of Best Fit"

\* regression line

### Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, $x$	0	1	2	3	3	5	5	5	6	7	7	10
Test score, $y$	96	85	82	74	95	68	76	84	58	65	75	50

- a. Sketch a graph of the distribution and then describe the distribution.



- b. Find the regression line.
- c. Find the correlation coefficient.
- d. Use the regression line to predict the test scores given the time online:

$x = 4$  hours

$x = 9$  hours

$x = 15$  hours

LinReg  
 $y = ax + b$   
 $a = -4.06741573$   
 $b = 93.97003745$   
 ~~$r = -0.8312962309$~~   
 $r = -0.8312962309$

Correlation coefficient

$r = -0.83$

→ Regression Line  
 $y = ax + b$   
 $y = -4.067x + 93.97$   
 $x = 4$   
 $y = -4.067(4) + 93.97$   
 $y = 77.703$   
 $x = 9$   
 $y = -4.067(9) + 93.97$   
 $y = 57.43$   
 $x = 15$   
 $y = -4.067(15) + 93.97$   
 $y = 33.07$

The image shows a TI-84 Plus calculator screen with the following text: `LinReg`, `y=ax+b`, `a=-4.06741573`, `b=93.97003745`, ~~`r=.6510534235`~~, and `r=-.8312962309`. A blue oval highlights the first four lines. To the right, blue handwritten text reads "Equation" and "Linear Regression". A red arrow points to the `r` value, with red handwritten text below it reading "correlation coefficient". Below the screen, blue handwritten text shows  $b) y = -4.067x + 93.97$  and red handwritten text shows  $c.) r = .83$ .

LinReg  
y=ax+b  
a=-4.06741573  
b=93.97003745  
~~r=-.8312962309~~  
r=-.8312962309  
correlation  
coefficient

$$y = ax + b$$
$$y = -4.067x + 93.97$$

x=4 → y=

r = - .83      x = 9      y =

x = 15      y =

```
LinReg  
y=ax+b  
a=-4.06741573  
b=93.97003745  
r2=.6910534235  
r=-.8312962309
```

## Notes 9.1 & 9.2 Correlation & Linear Regression

### Correlation

- A correlation is a **comparison between 2 quantities**
- The data can be represented by  $(x, y)$  where  $x$  is the **independent** or **input** variable, and  $y$  is the **dependent** or **output** variable. **domain**
- One way to determine whether a linear correlation exists between two variables is to use a **range**.

\* Regression Line

Linear Regression "Test"  
"the line of best fit"

### Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, $x$	0	1	2	3	3	5	5	5	6	7	7	10
Test score, $y$	96	85	82	74	95	68	76	84	58	65	75	50

- a. Sketch a graph of the distribution and then describe the distribution.



- b. Find the regression line.
- c. Find the correlation coefficient.
- d. Use the regression line to predict the test scores given the time online:

$$x = 4 \text{ hours}$$

$$x = 9 \text{ hours}$$

$$x = 15 \text{ hours}$$

LinReg

$$y = ax + b$$

$$a = -4.06741573$$

$$b = 93.97003745$$

~~$$r = -0.8312962309$$~~

~~$$r = -0.8312962309$$~~

$$r = -0.8312962309$$

correlation coefficient

$$r = -0.83$$

$$x = 4$$

$$y = -4.067(4) + 93.97$$

$$y = 77.7$$

$$y = -4.067(9) + 93.97$$

$$y = 57.37$$

$$y = -4.067(15) + 93.97$$

$$y = 32.97$$

$$y = mx + b$$

$$y = ax + b$$

Regression Line

$$y = -4.067 + 93.97$$

## Notes 9.1 & 9.2 Correlation & Linear Regression

### Example 2:

The budgets & worldwide grosses of 15 of the most expensive 20<sup>th</sup> Century Fox Movies are shown.

<b>Budget, x</b> (millions)	200	150	125	125	115	115	115	110	110	110	105	102	100	100	100
<b>Gross, y</b> (millions)	1835.4	459.4	406.4	542.7	924.3	656.7	848.5	571.1	211.4	150.5	348.8	358.8	365.3	359.1	249.0

- a. Sketch a graph of the distribution and then describe the distribution.



- b. Find the regression line.
- c. Find the correlation coefficient.
- d. Use the regression line to predict the gross amount of money for the given budget:

*\$120,000,000*

*\$93,000,000*

2nd + Catalog then... scroll down to DiagnosticON

Press Enter to get the words "Done"

Input x into L1      Input y into L2

To get Equation & "r" Stat + scroll over to Calc #4 LinReg (ax + b)

2 million = \$ 2,000,000

2.4 million = \$ 2,400,000

x = 120,000,000

120 million

