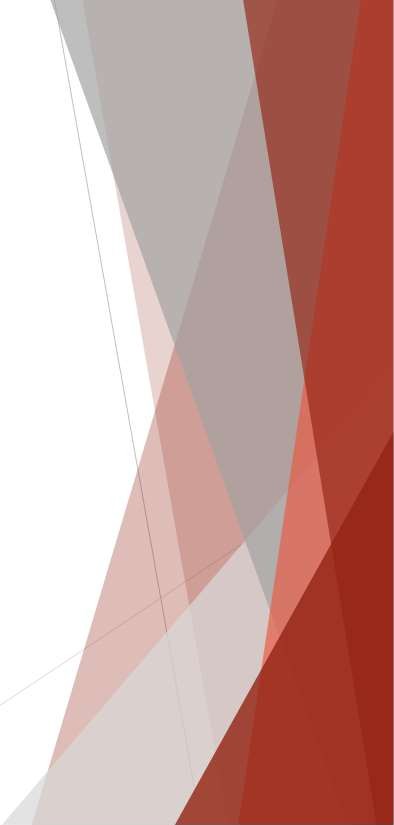


Business Analytics

Chapter 7 Linear Regression

4



Introduction

 Managerial decisions based on:

 Relationship between two or more variables:

 Example:

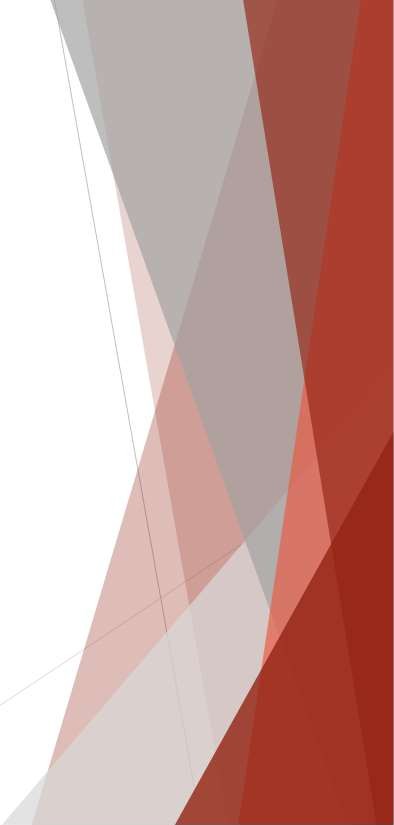
 **Paying for ads** vs. **getting sales**

 Intuition can be useful but…

 If you have data

 **USE regression analysis** to show how the variables are related.

5

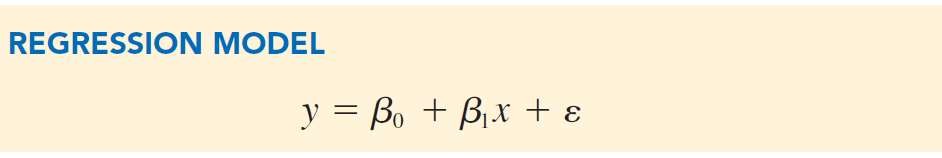
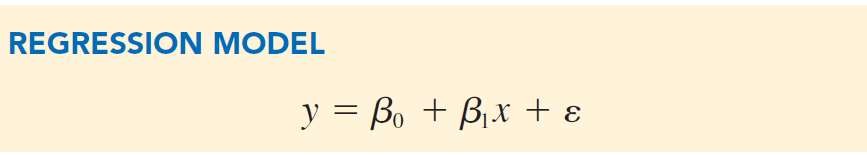
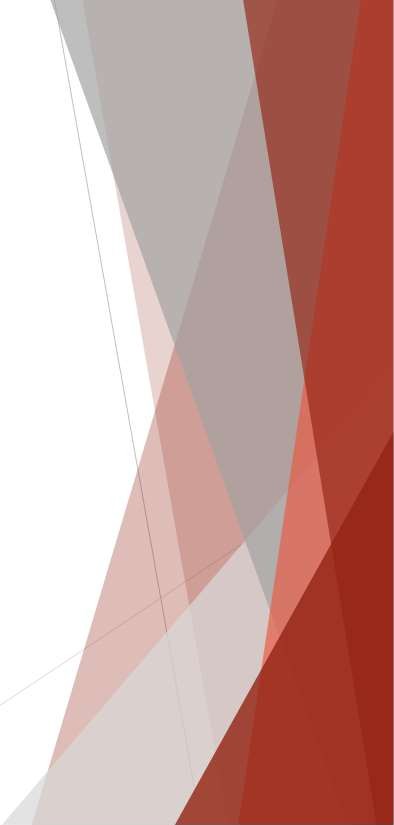


Simple Linear Regression Model

Regression Model

Estimated Regression Equation

6



Simple Linear Regression Model

Best Fit Linear Regression

Slope intercept Form

𝑦 =𝑦𝑚𝑥 + 𝑏

**Best Fit Regression Model:**

 The equation that describes how y is related to x with an intercept.

 Slope = 𝛽1

 Intercept = 𝛽0

 Y = dependent variable

 X = independent variable

**Equation of a line:**



The equation that describes how y is related to x with an intercept

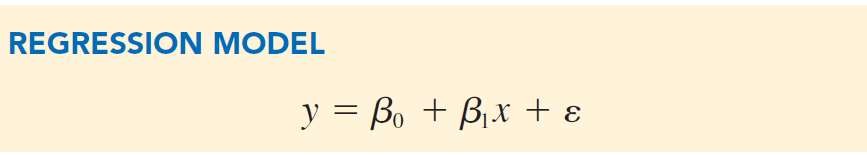
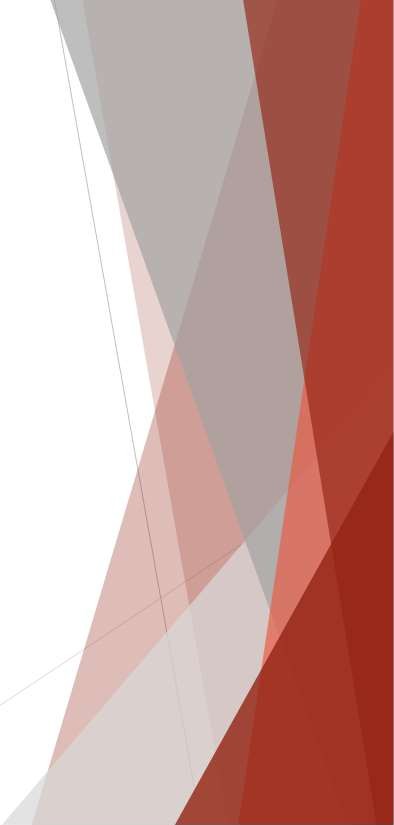
 Slope = m

 Intercept = b

 Y = dependent variable

 X = independent variable

7



Simple Linear Regression Model

 The equation that describes the “True Relationship” between y and x and an error term.

**Simple Linear Regression Model:**

 Parameters: The characteristics of the population, 𝛽0 and 𝛽1.

 Slope = 𝛽1

 Intercept = 𝛽0

 𝜀 = error term

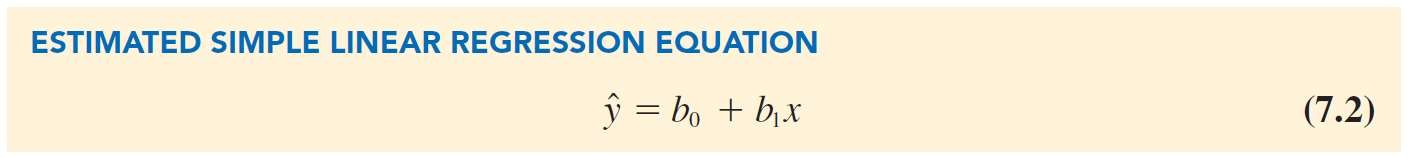
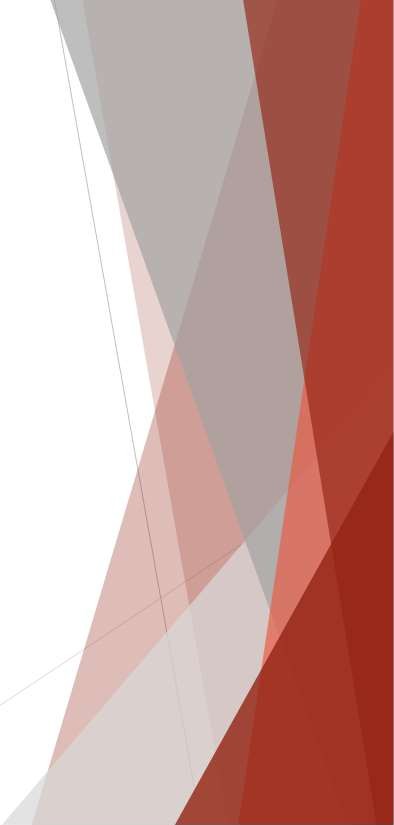
**Example:**

**How is travel time (y) of a delivery truck related to number miles traveled (x)**

 Variability in y that cannot be explained by the relationship between x and y

 Assume 𝜀 is normally distributed with mean 0 and constant variance

8



Simple Linear Regression Model

**Estimated Regression Equation:**

 The “true” parameter values are usually not known

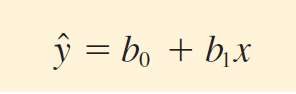
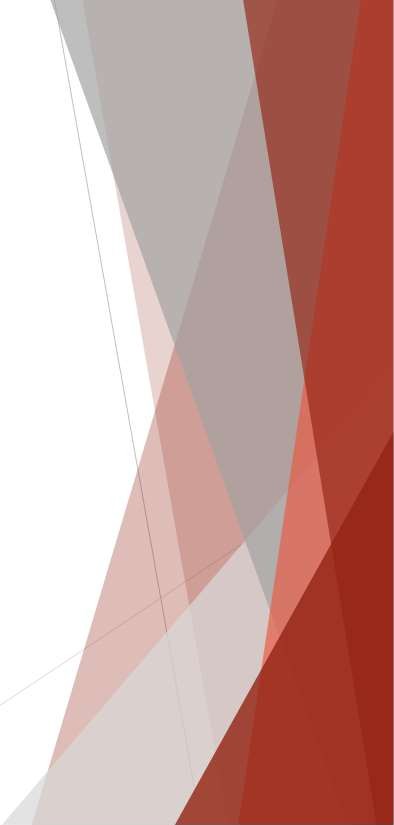
 must be estimated using sample data.

 Sample Statistics - 𝑏0 and 𝑏1 are calculated as estimates of 𝛽0 and 𝛽1

 We plug in 𝑏0 and 𝑏1 and drop the error term

 Expected value of 𝜀 is = 0

9



Simple Linear Regression Model

In the estimated simple linear regression equation:

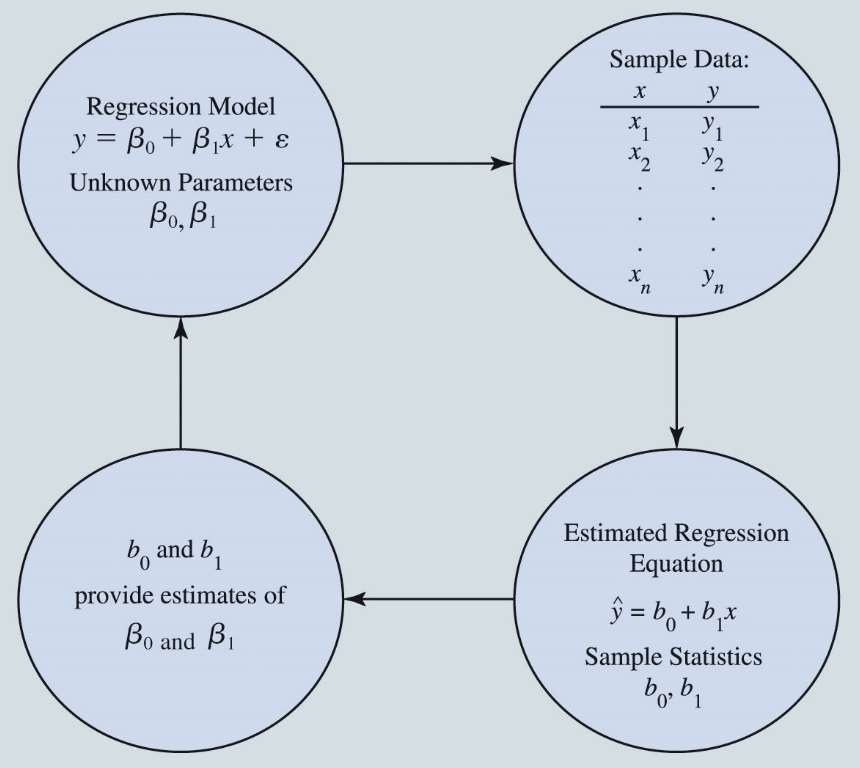
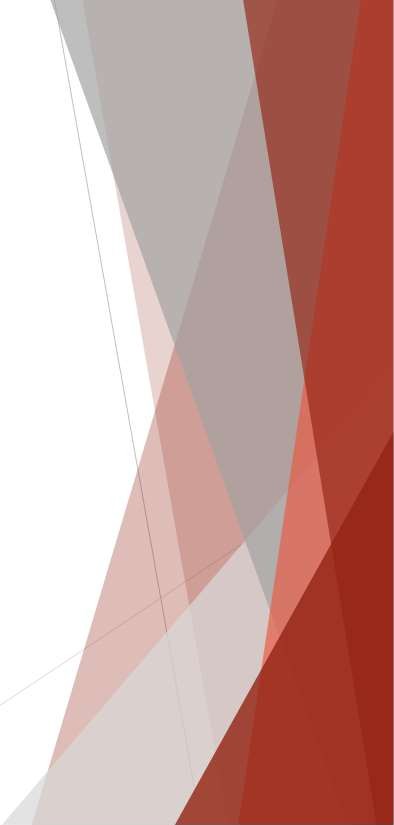
*y*ˆ = Estimate for the mean value of *y* corresponding to a given value of *x*. *b*0 = Estimated *y*-intercept.

*b*1 = Estimated slope.

 The graph of the estimated simple linear regression equation is called the estimated regression line*.*

 "In general, 𝑦^ is the point estimator of 𝐸 𝑦| 𝑥 , the mean value of y for a given value of x

10



**Example**

Simple Linear Regression Model

The Estimation Process in Simple Linear

Regression

 Butler Trucking

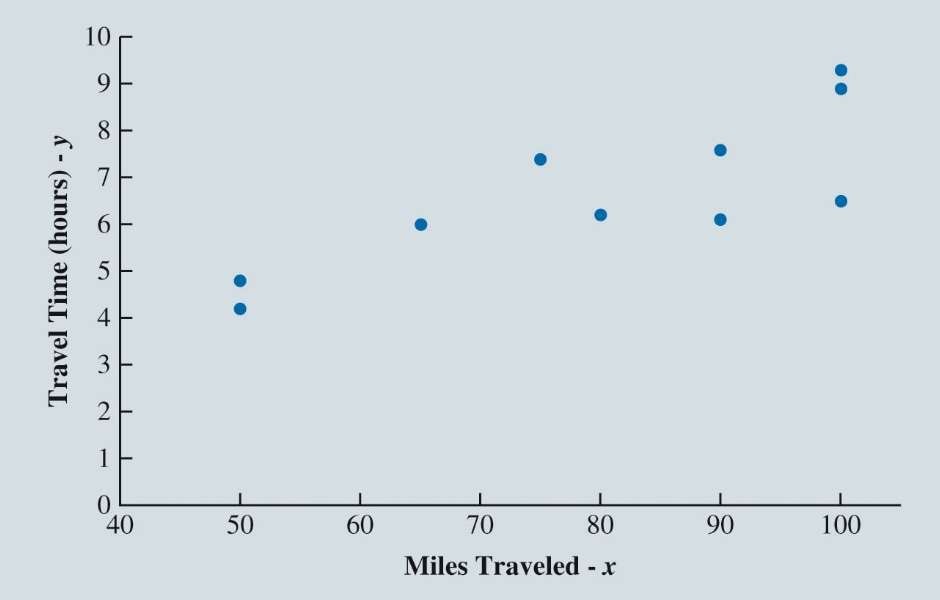
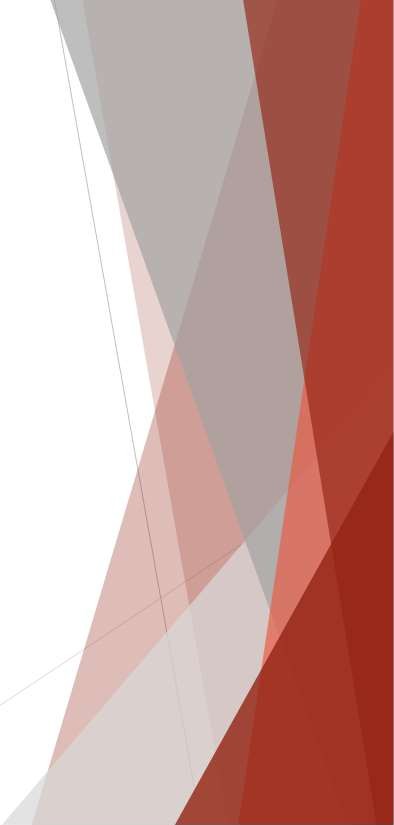
 How is travel time (y) of a delivery truck related to number miles traveled (x)

 We need data….

11

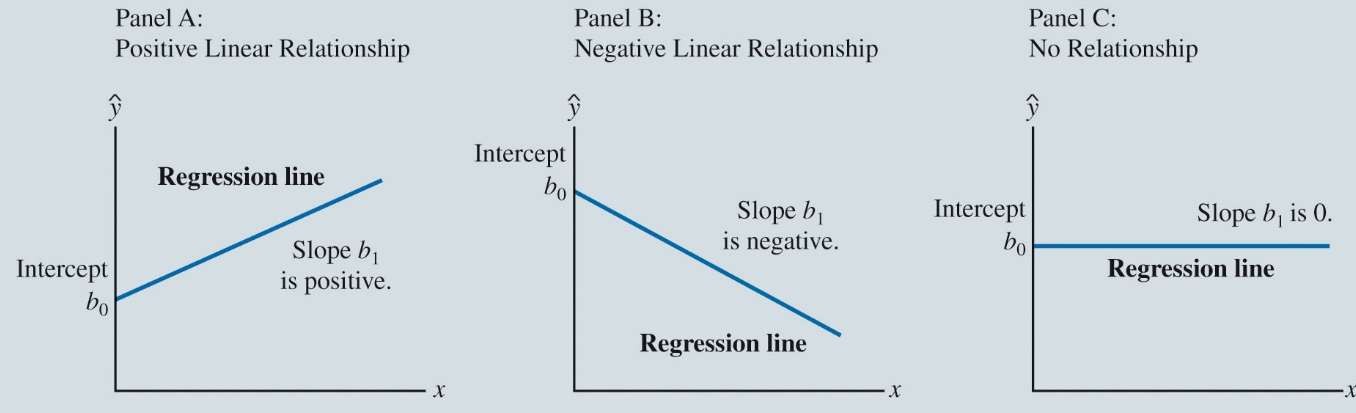
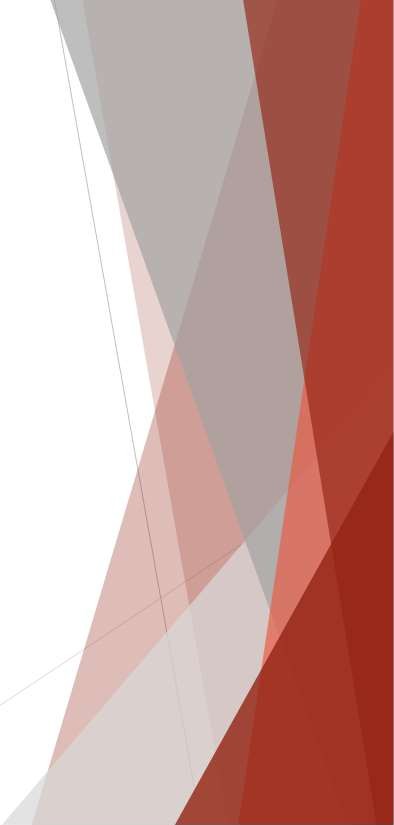
|  |  |  |
| --- | --- | --- |
| **Driving**  **Assignment *i*** | ***x* = Miles**  **Traveled** | ***y =* Travel Time**  **(hours)** |
| 1 | 100 | 9.3 |
| 2 | 50 | 4.8 |
| 3 | 50 | 8.9 |
| 4 | 100 | 6.5 |
| 5 | 50 | 4.2 |
| 6 | 80 | 6.2 |
| 7 | 75 | 7.4 |
| 8 | 65 | 6.0 |
| 9 | 90 | 7.6 |
| 10 | 90 | 6.1 |

12



Least Squares Method

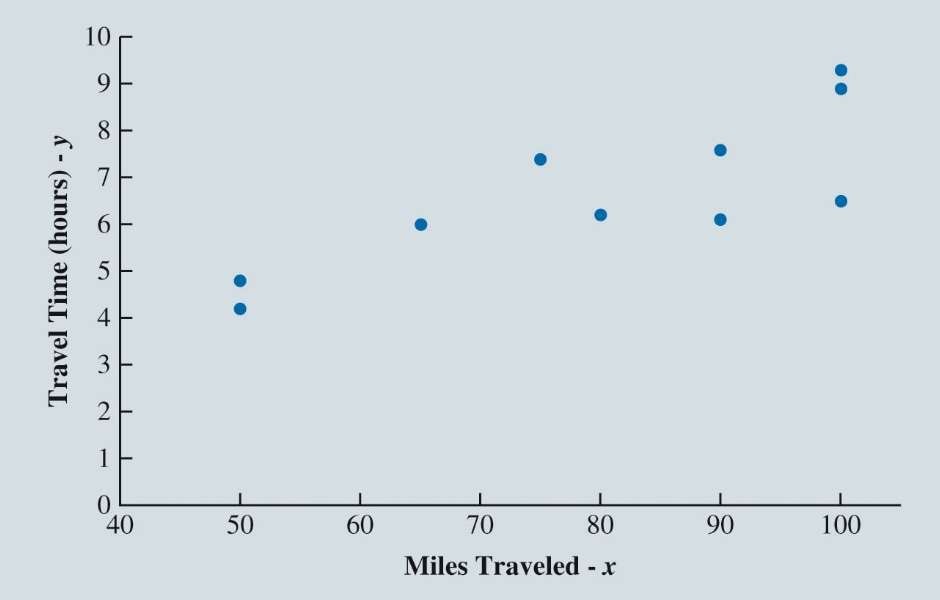
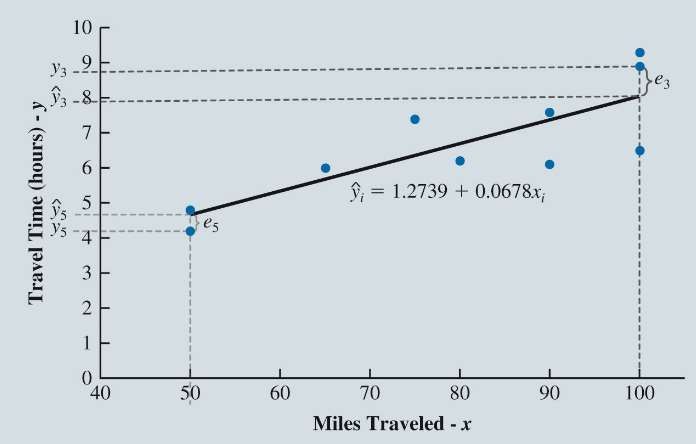
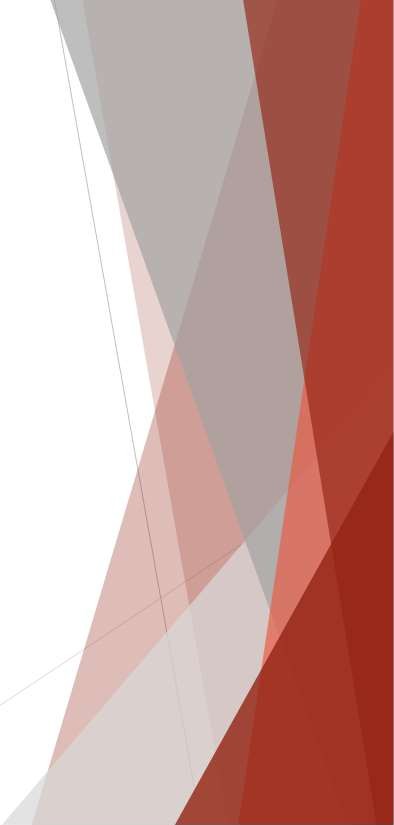
Miles Traveled and Travel Time for 10 Butler Trucking Company Driving Assignments



Simple Linear Regression Model

Figure 7.2: Possible Regression Lines in Simple Linear Regression

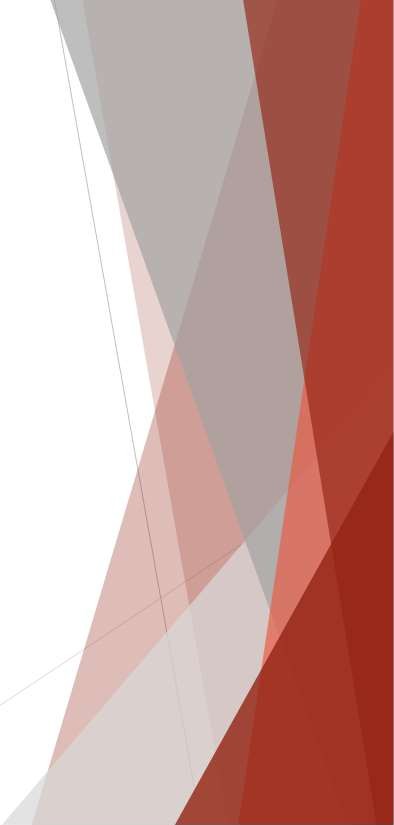
13



Least Squares Method

Scatter Chart of Miles Traveled and Travel Time for Butler Trucking Company Driving Assignments with Regression Line Superimposed

14



Least Squares Method

**Least Squares Estimates of the Regression Parameters:**

 For the Butler Trucking Company data

 Estimated slope of 𝑏1 = 0.0678

 Estimated y-intercept of 𝑏0= 1.2739

The estimated simple linear regression model:

*y*ˆ = 1.2739 + 0.0678*x*1

**What do these numbers mean?**

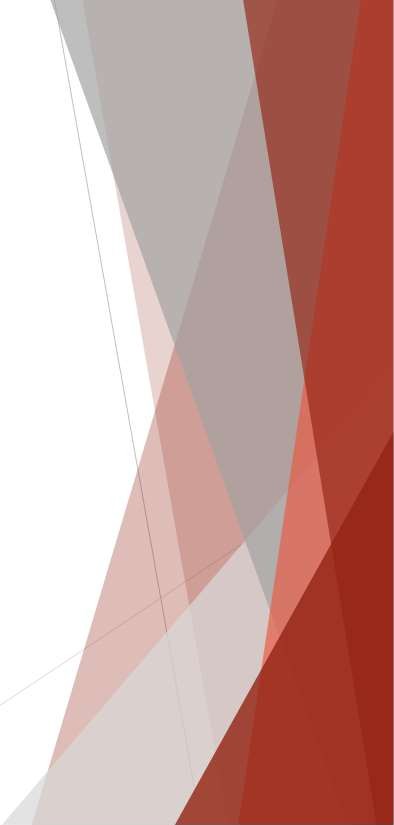
 𝑏1 = 0.0678 – As the trip increases 1 more mile,

 the average travel time increases 0.0678 hours (4 Minutes)

 𝑏0= 1.2739 – When the trip is 0 miles,

 the estimated travel time is 1.2 hours (76 minutes)

15



Examples

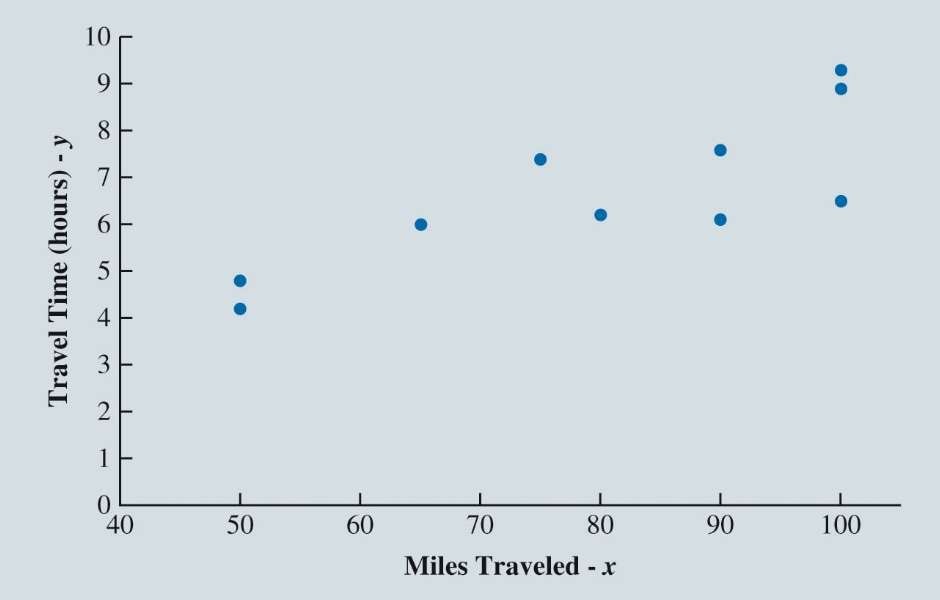
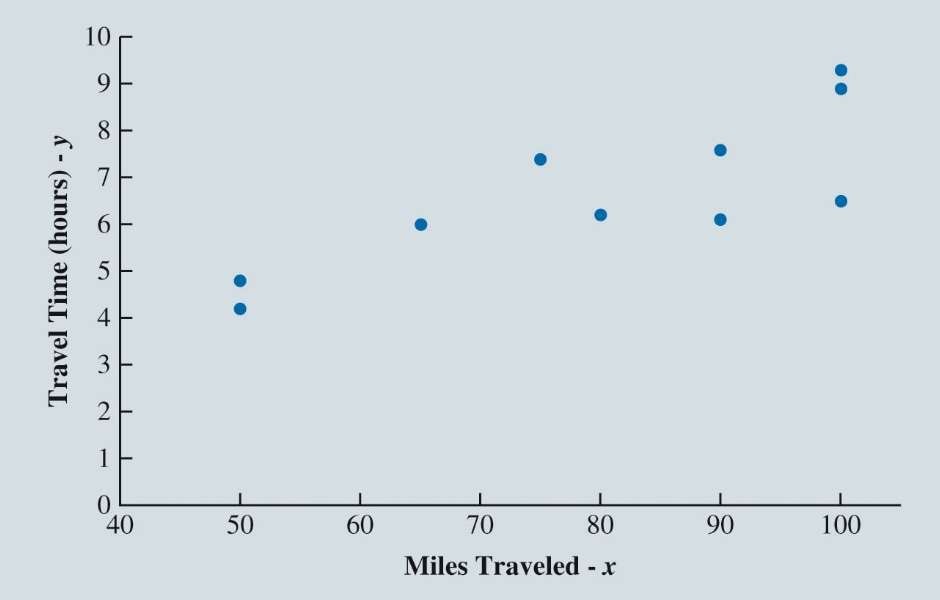
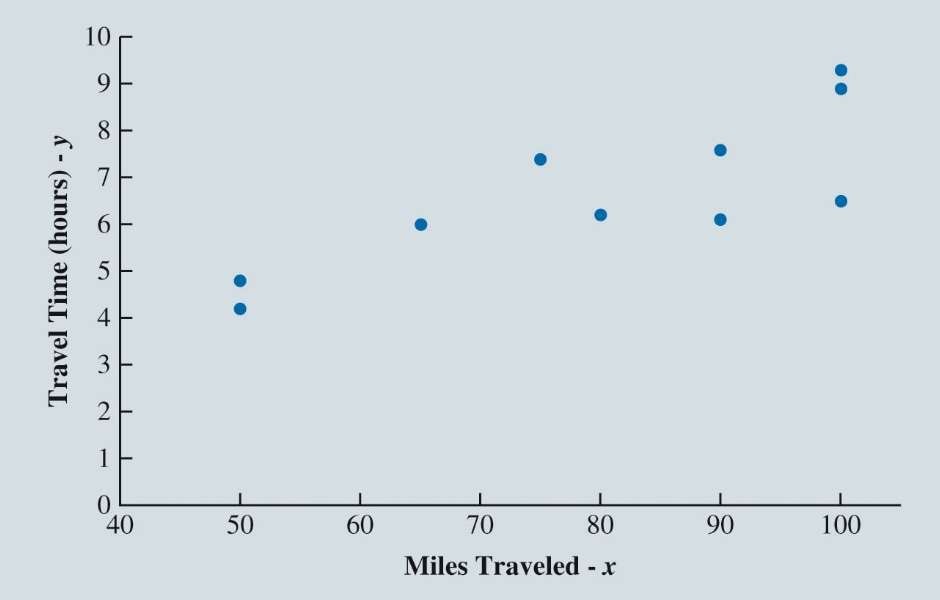
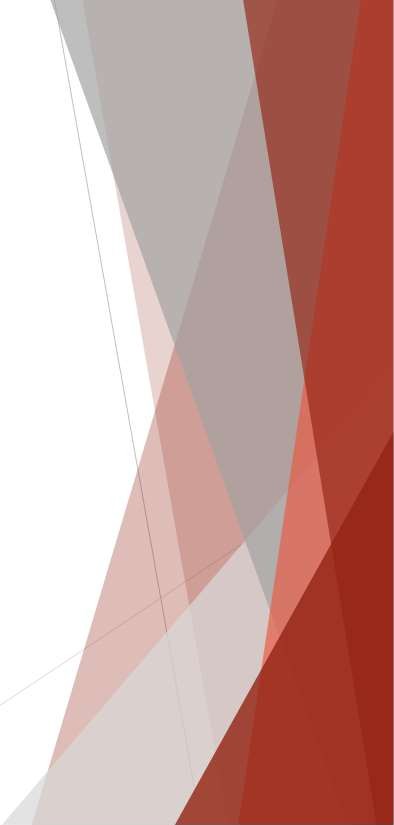
**What is the expected travel time (y) of a delivery truck that travels 75 miles on deliveries?**

= 1.2739 + 0.0678 (75) = 6.35 hours

**What is the expected travel time (y) of a delivery truck that travels 100 miles on deliveries?**

= 1.2739 + 0.0678 (100) = 8.05 hours

16



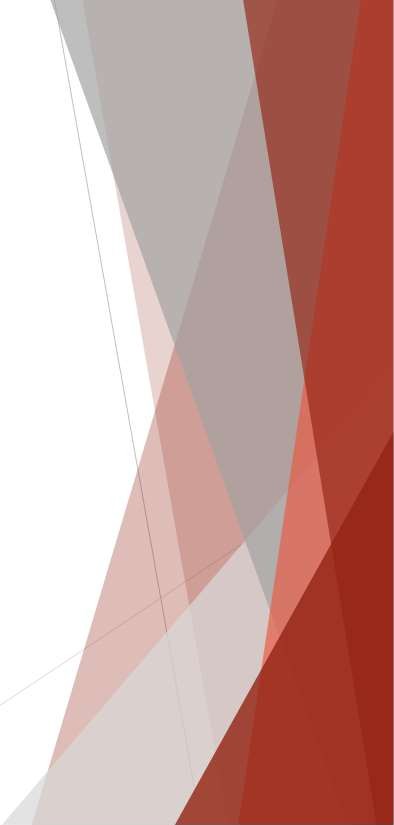
How do we know what line is best?

A.

B.

C.

17

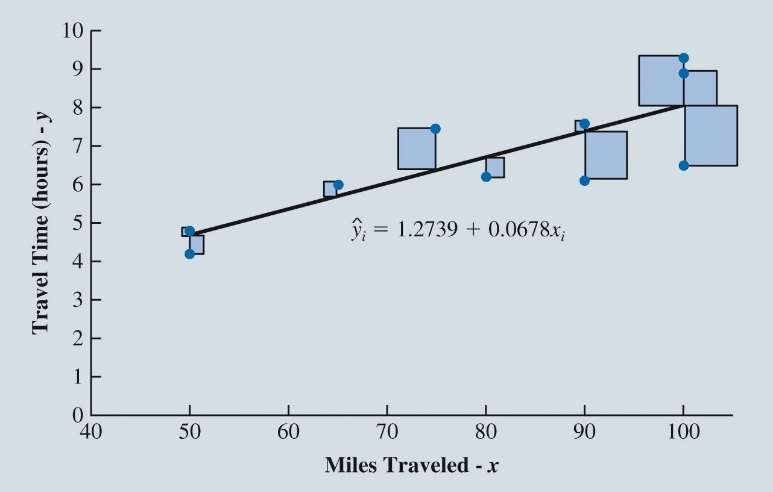
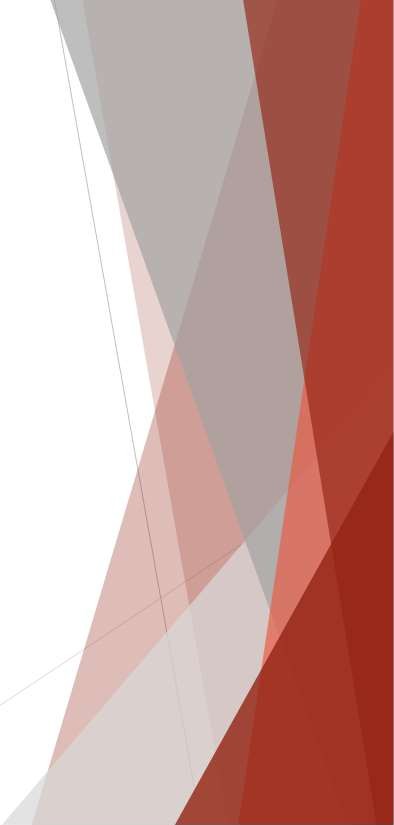


Least Squares Method

Least Squares Estimates of the Regression Parameters

Using Excel’s Chart Tools to Compute the Estimated Regression Equation

18



Least Squares Method

**Least Squares Method**

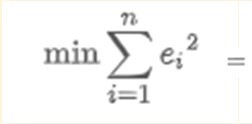
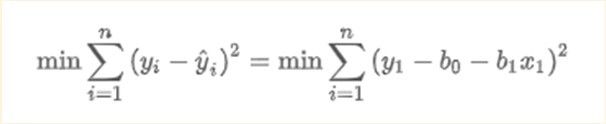
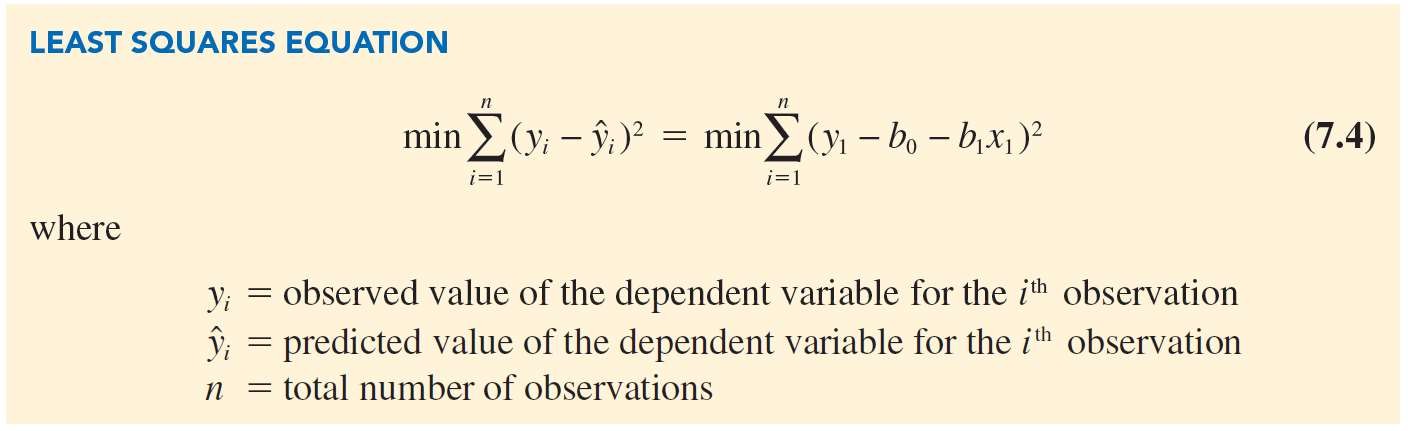
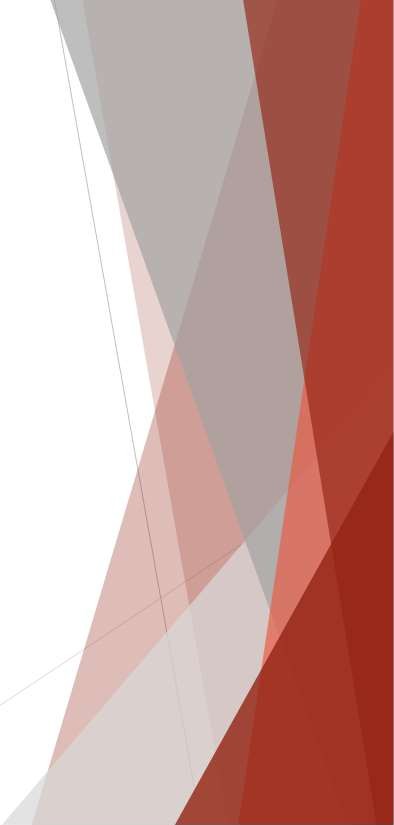
1. Measure the difference between each y value (𝑦i data point) and the estimated y value on the regression line (𝑦^i)

Denoted as 𝑒i = 𝑦i − 𝑦^i

(Called Residual)

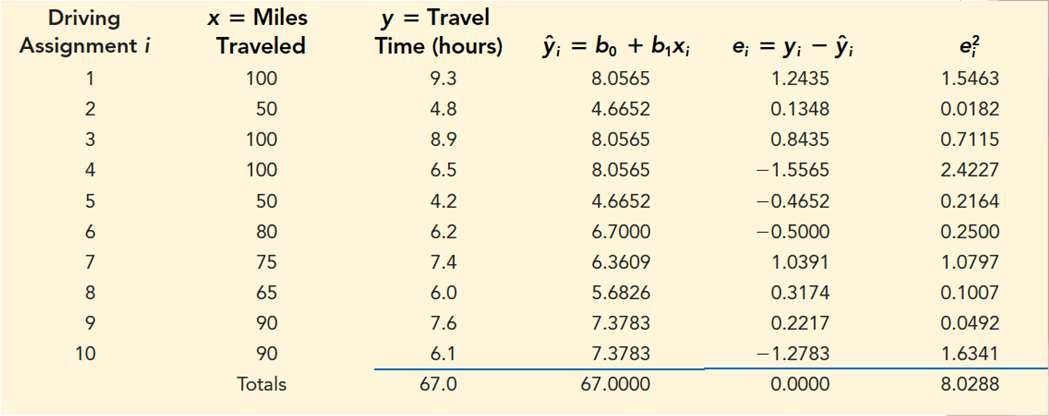
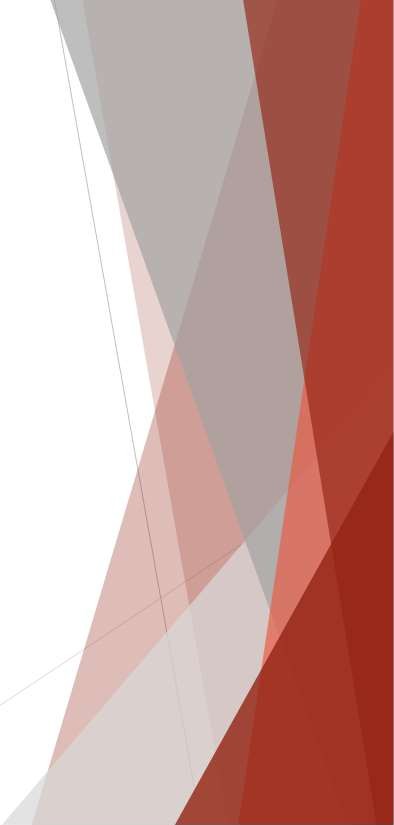
1. Square the differences
2. Add up the squared differences
3. The minimum sum of square differences is the best-fit line

19



Least Squares Method

20

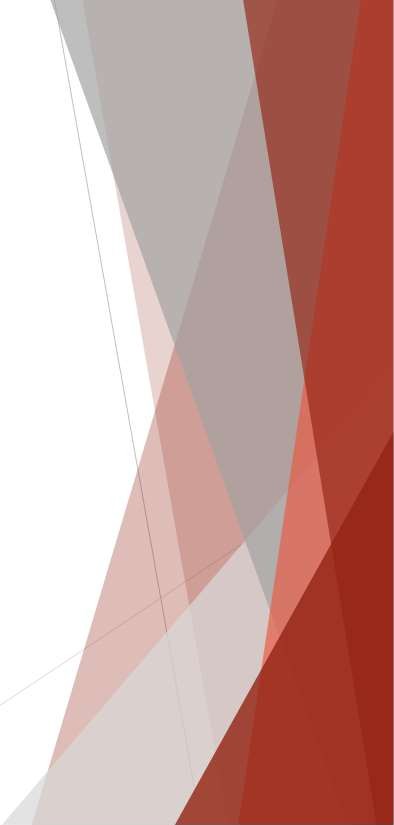


Least Squares Method

Table 7.2: Predicted Travel Time and Residuals for 10 Butler Trucking Company Driving Assignments

What do you notice about the actual y values (data points) and the estimated y values (𝑦^i) What do you notice about the sum of residuals or error terms?

21



*i*

*yi* = value of the dependent variable for the *i*th observation.

*x* = mean value for the independent variable.

*y* = mean value for the dependent variable.

*n* = total number of observations.

*i* =1

*x* = value of the independent variable for the *i*th observation.

2

)

*x* – *x*

1

0

*b* = *y* – *b x*

(

*n*

Σ *i*

1

Σ( *xi* – *x* )(*yi* – *y* )

*b* = *i* =1

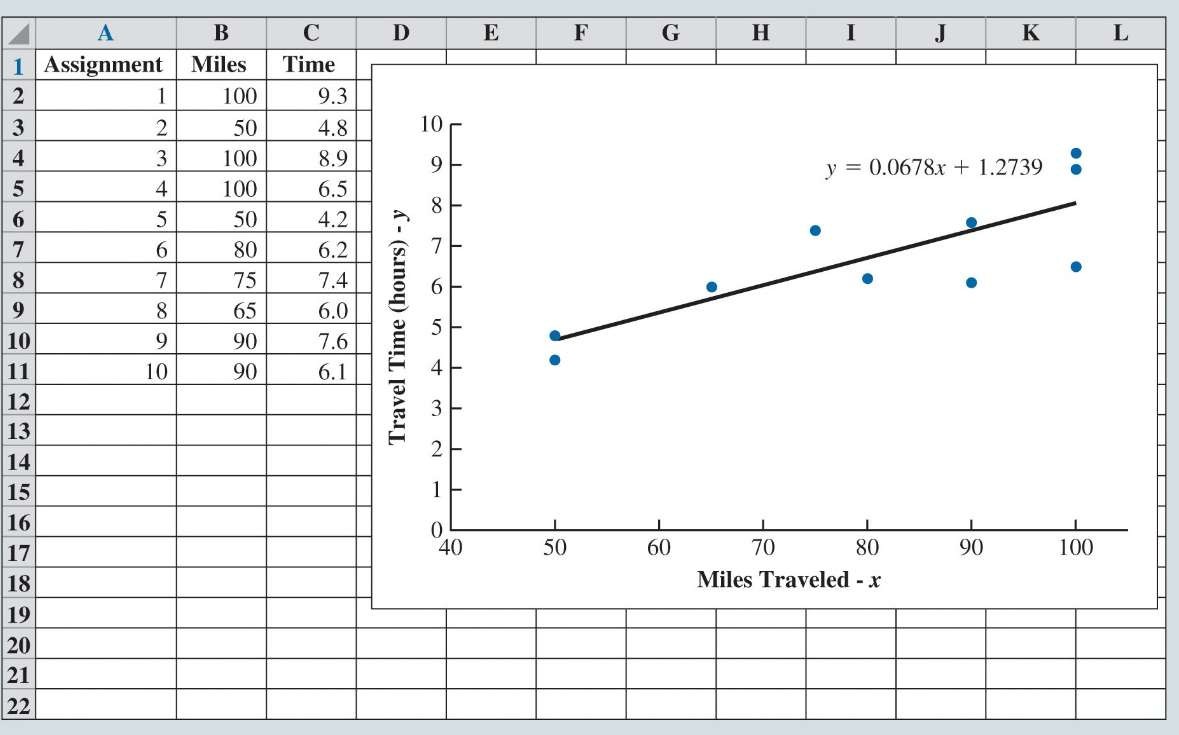
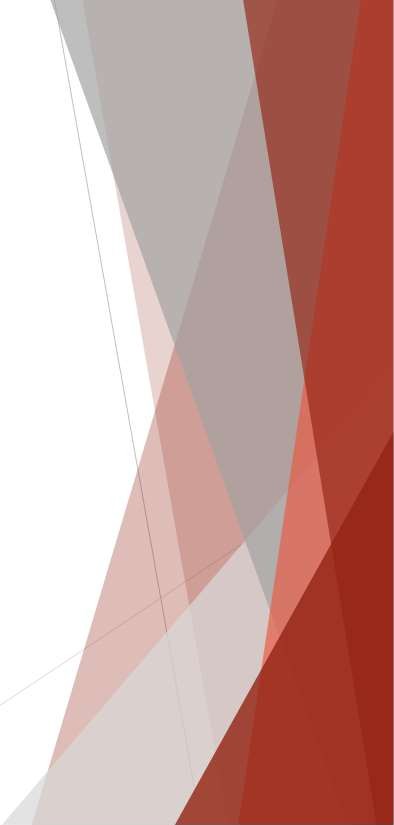
*n*

*y*-Intercept Equation

Slope Equation

Least Squares Method – Differential Calculus

22



Least Squares Method

**Using Excel’s Chart Tools to Compute the Estimated Regression Equation:**

 After constructing a scatter chart with Excel’s chart tools:

1. Right-click on any data point and select **Add Trendline.**

2. When the **Format Trendline** task pane appears:

 Select **Linear** in the **Trendline Options**

area.

 Select **Display Equation on chart** in the

**Trendline Options** area.

23