FROM BEFORE

Linear Models

Regression Models

X Variables are Interval or Ratio Scaled

Mean Structure Models

(Analysis of Variance Models)

X Variables are Nominal or Ordinal Scaled

Linear Models

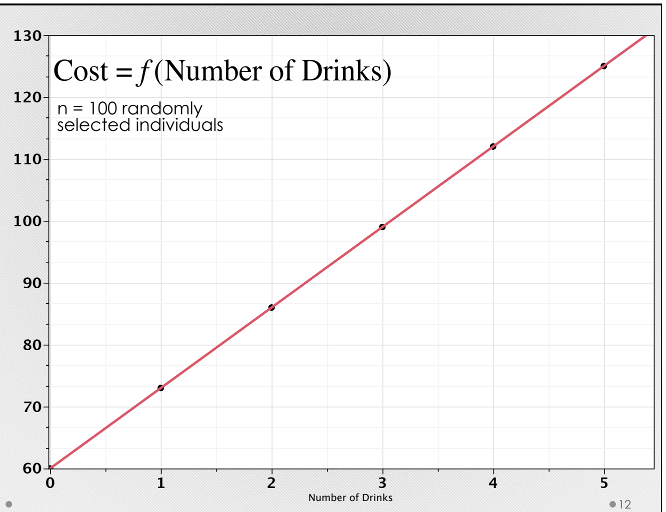
Regression Models

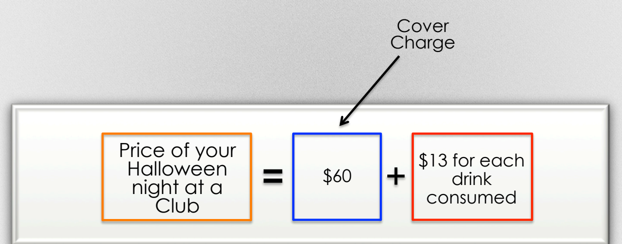
X Variables are Interval or Ratio Scaled

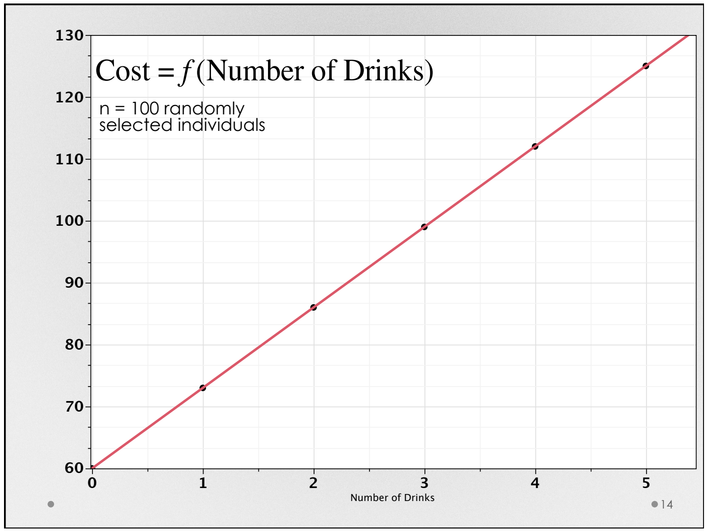
Mean Structure Models

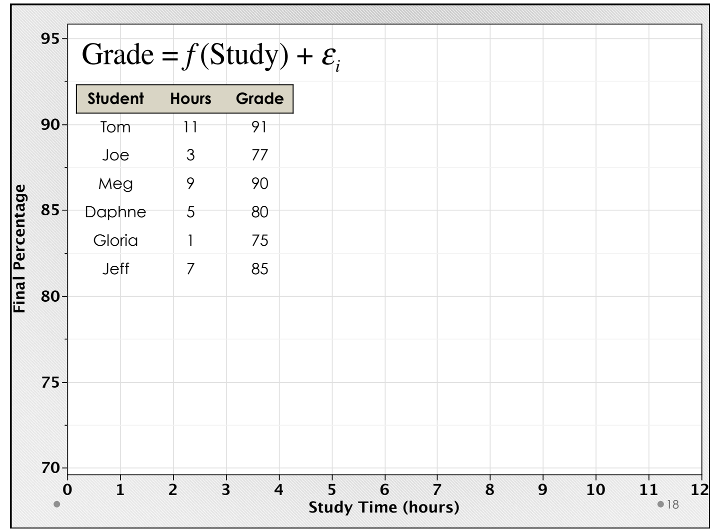
(Analysis of Variance Models)

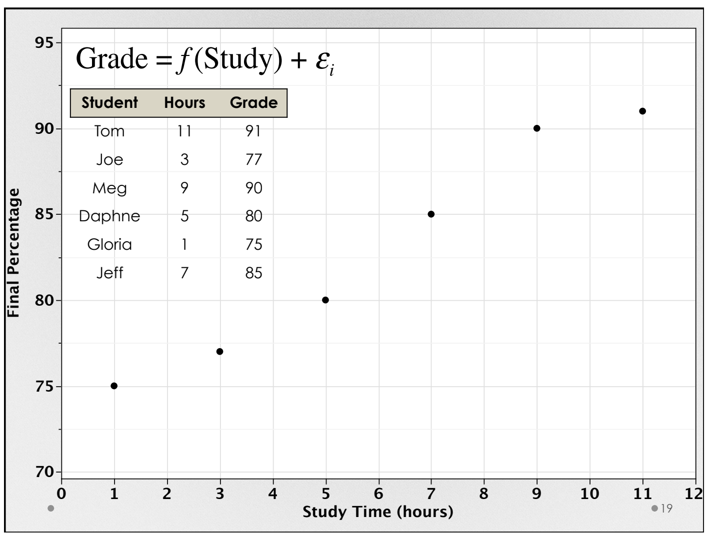
X Variables are Nominal or Ordinal Scaled

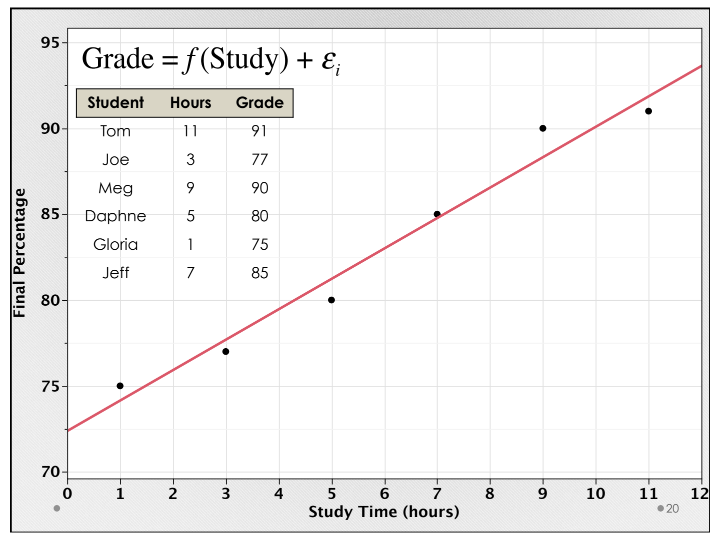





Cost = f(Number of Drinks)
n = 100 randomly selected individuals
130
120
110
100
90
80
70
60
0
1
2
3
4
5
Number of Drinks
14


Grade = f(Study) + ε_i
Student Hours Grade
Tom 11 91
Joe 3 77
Meg 9 90
Daphne 5 80
Gloria 1 75
Jeff 7 85
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
18


Grade = f(Study) + ε_i
Student Hours Grade
Tom 11 91
Joe 3 77
Meg 9 90
Daphne 5 80
Gloria 1 75
Jeff 7 85
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
19


Grade = f(Study) + ε_i
Student Hours Grade
Tom 11 91
Joe 3 77
Meg 9 90
Daphne 5 80
Gloria 1 75
Jeff 7 85
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
20

The Regression Line

Regression

Regression is a statistical technique for finding the best-fitting "line" for a set of data.

The result is called a regression line.

Simple Linear Regression

Simple linear regression is a procedure for finding the best fitting straight line for a set of data using a single predictor variable.

The result is called the linear regression line.

One Factor Linear Model
(Population Model)

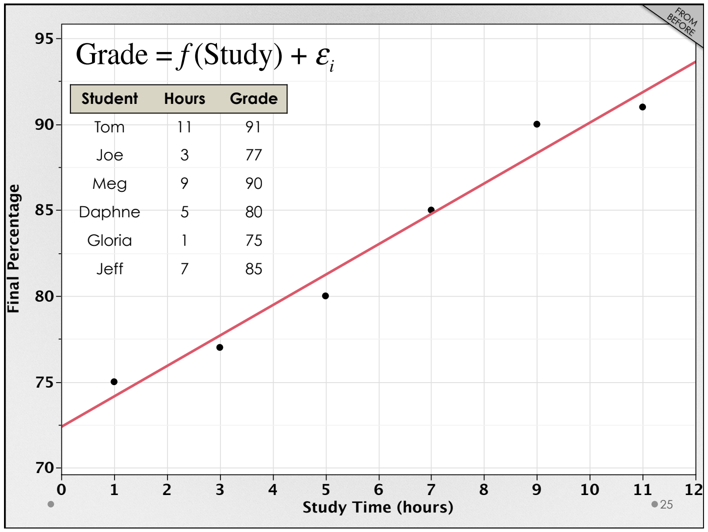
FROM BEFORE

$$\boxed{Y_{ij}} = \boxed{\mu_{..}} + \boxed{\tau_j} + \boxed{\varepsilon_{ij}}$$

Score on Y for the ith individual in the jth group = Grand Mean + Treatment Effect for Group j + Error

$$\text{Grade} = f(\text{Study}) + \varepsilon_i$$

Student	Hours	Grade
Tom	11	91
Joe	3	77
Meg	9	90
Daphne	5	80
Gloria	1	75
Jeff	7	85



One Factor Linear Model
(Population Model)

$$\boxed{Y_{ij}} = \boxed{\mu_{..}} + \boxed{\tau_j} + \boxed{\varepsilon_{ij}}$$

Score on Y for the ith individual in the jth group = Grand Mean + Treatment Effect for Group j + Error

One Factor Linear Model
(Population Model)

$$\boxed{Y_{ij}} = \boxed{\mu_{..}} + \boxed{\tau_j} + \boxed{\varepsilon_{ij}}$$

Score on Y for the ith individual in the jth group = Grand Mean + Treatment Effect for Group j + Error

One Factor Linear Model
(Population Model)

$$\boxed{\tau_j}$$

Treatment Effect for Group j

One Predictor Linear Regression Model
(Population Model)

$$\boxed{X_i}$$

Score on X for the ith individual

One Predictor Linear Regression Model
(Population Model)

$$\boxed{\beta_1 X_i}$$

Slope (Effect) × Score on X for the ith individual

One Predictor Linear Regression Model
(Population Model)

$$\boxed{\beta_0} + \boxed{\beta_1 X_i}$$

Y Intercept + Slope (Effect) × Score on X for the ith individual

One Predictor Linear Regression Model
(Population Model)

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1 X_i} + \boxed{\varepsilon_i}$$

Score on Y for the ith individual = Y Intercept + Slope (Effect) × Score on X for the ith individual + Error

One Predictor Linear Regression Model
(Population Model)

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i} + \boxed{\varepsilon_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

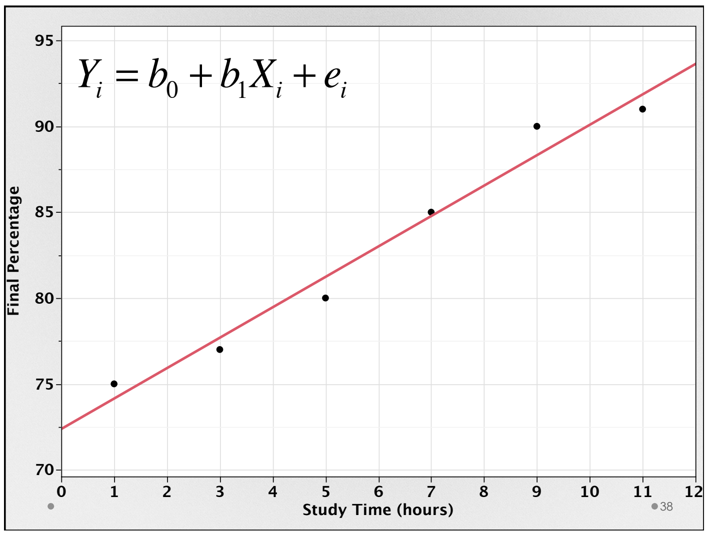
33

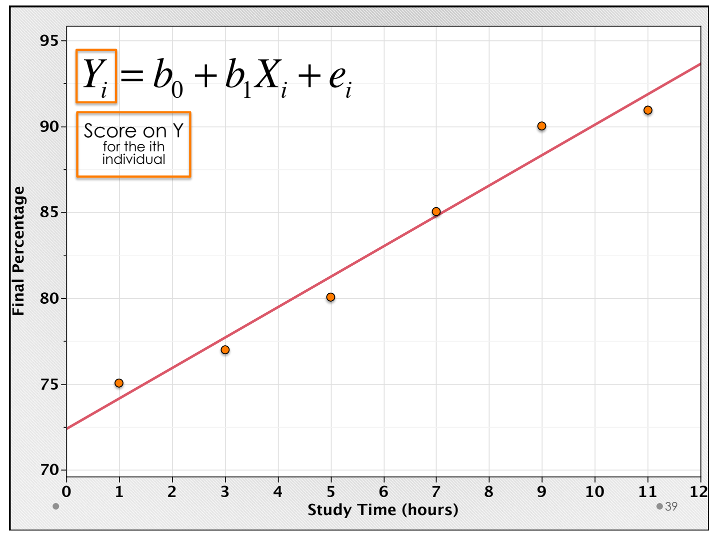
One Predictor Linear Regression Model
(Sample Model)

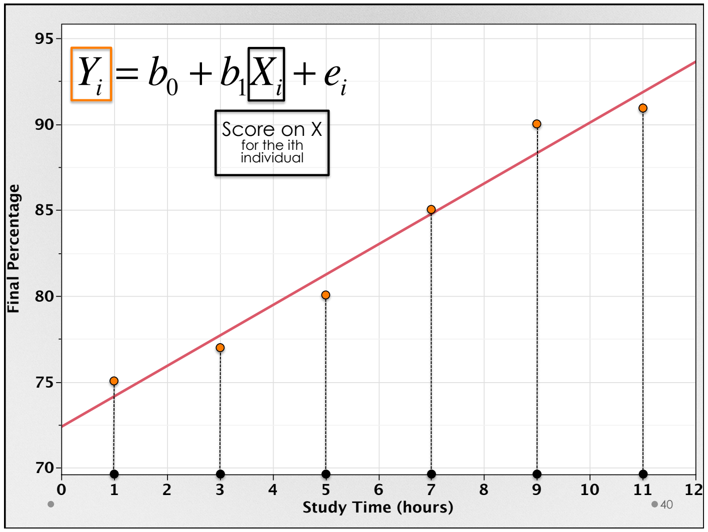
$$\boxed{Y_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i} + \boxed{e_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

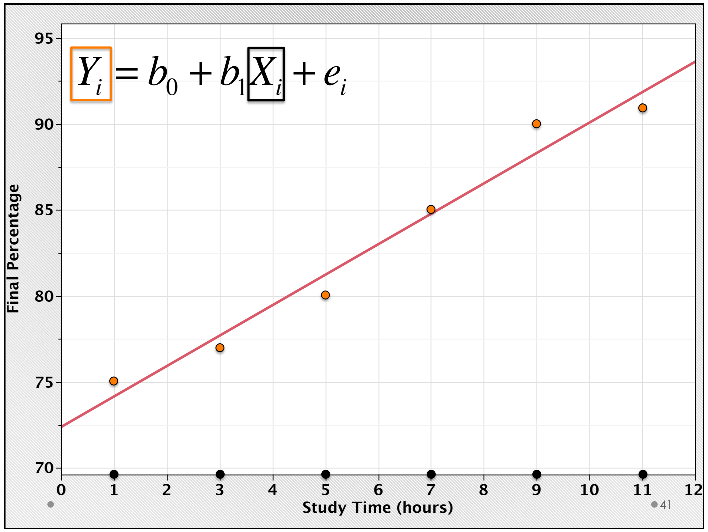
37



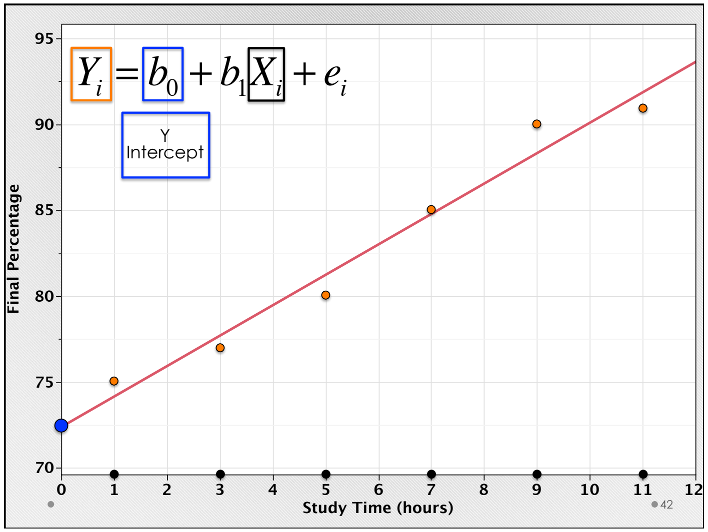




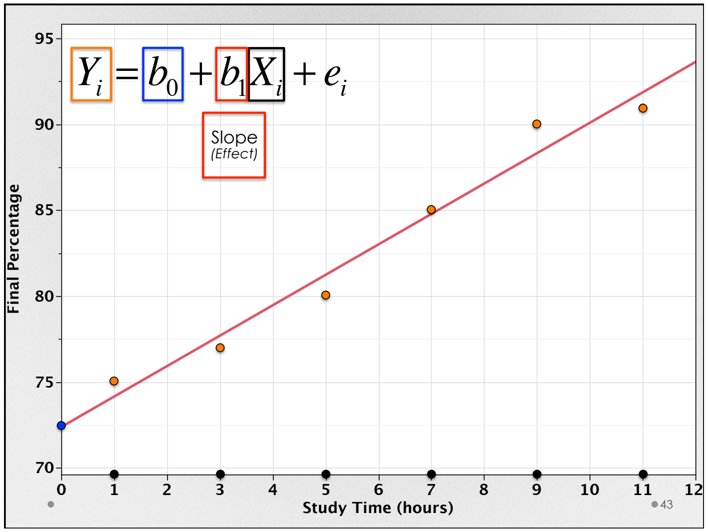
$Y_i = b_0 + b_1 X_i + e_i$
Score on X
for the ith
individual
Final Percentage
Study Time (hours)
40



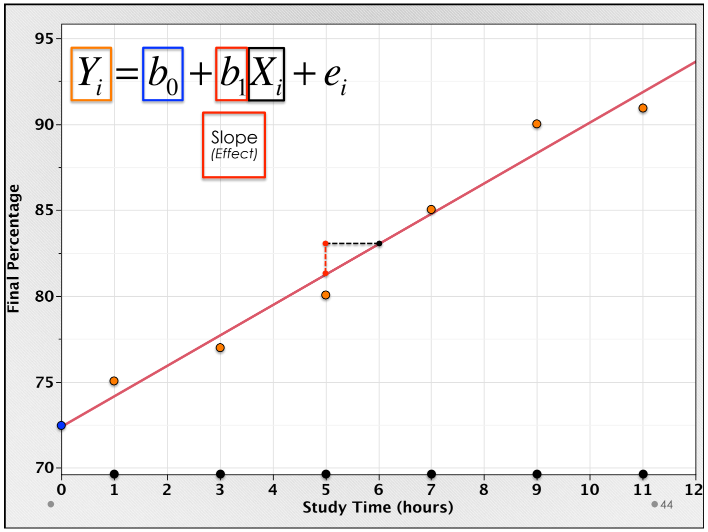
$Y_i = b_0 + b_1 X_i + e_i$
Final Percentage
Study Time (hours)
41

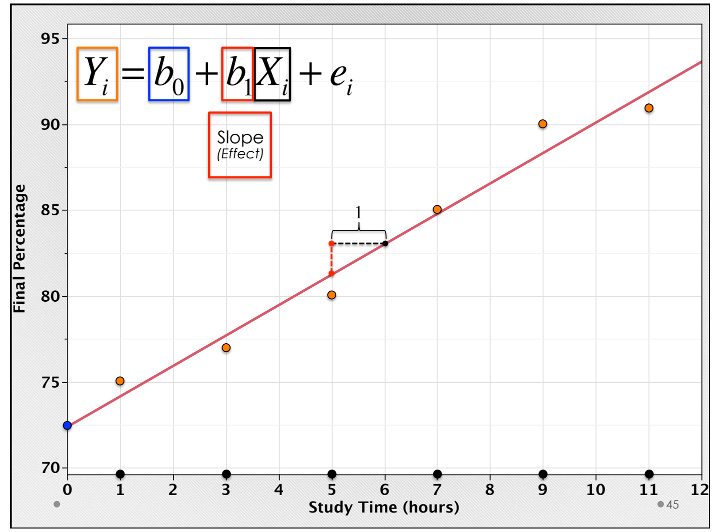


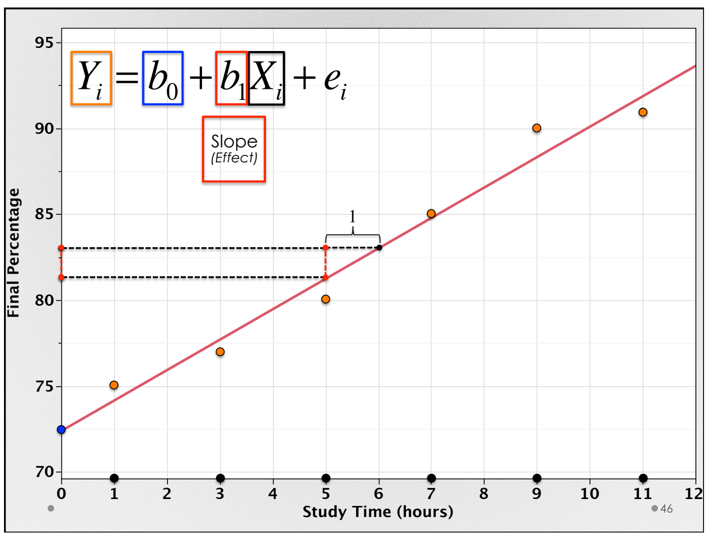
$Y_i = b_0 + b_1 X_i + e_i$
Y
Intercept
Final Percentage
Study Time (hours)
42

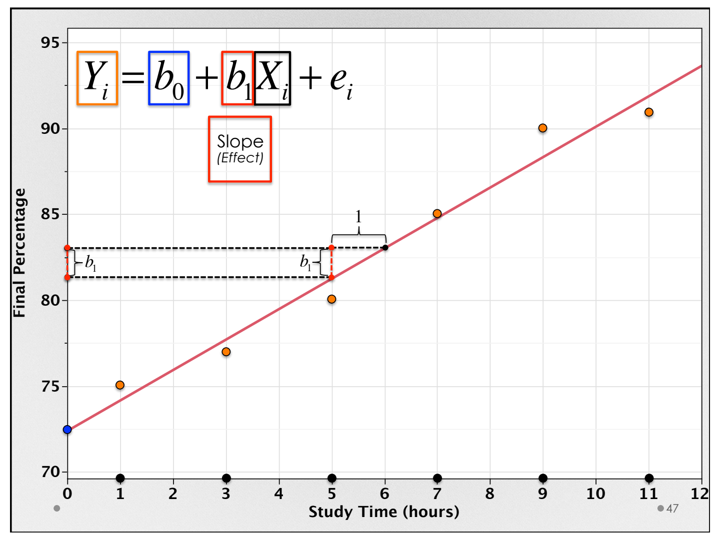


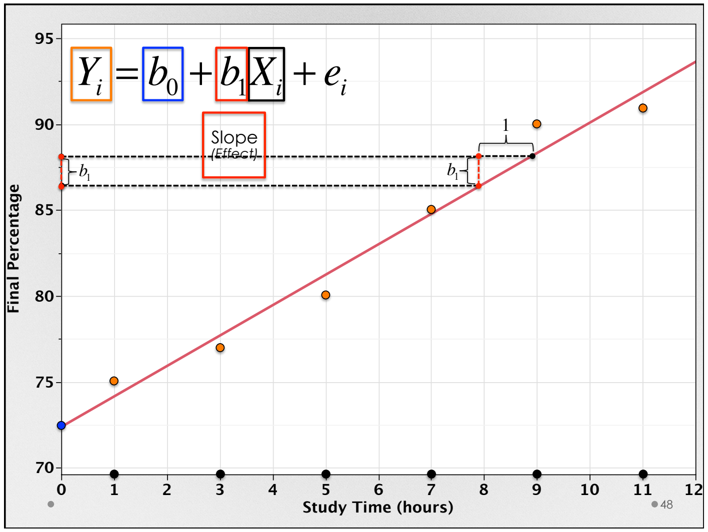
$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
Final Percentage
Study Time (hours)
43

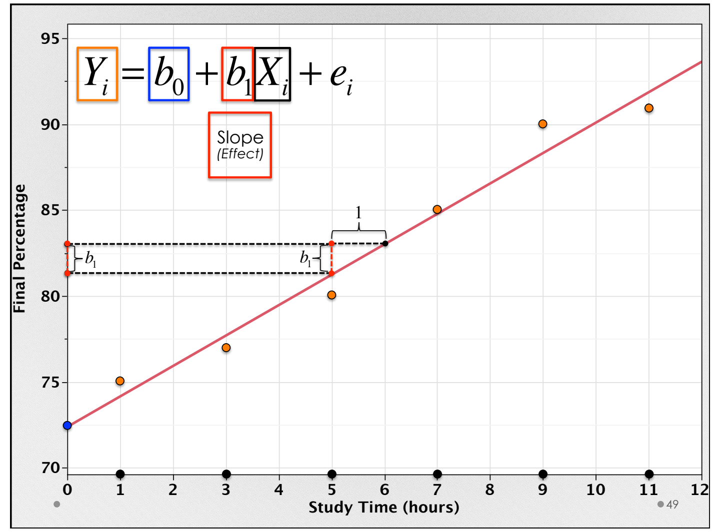

$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
44

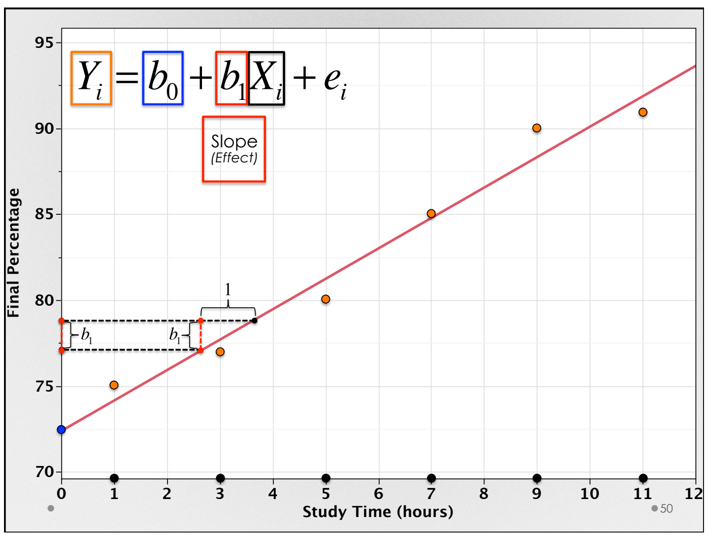

$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
45

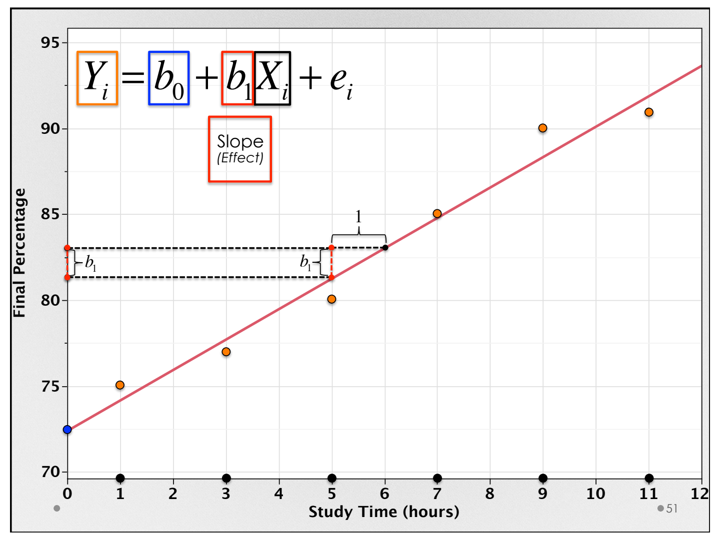

$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
46

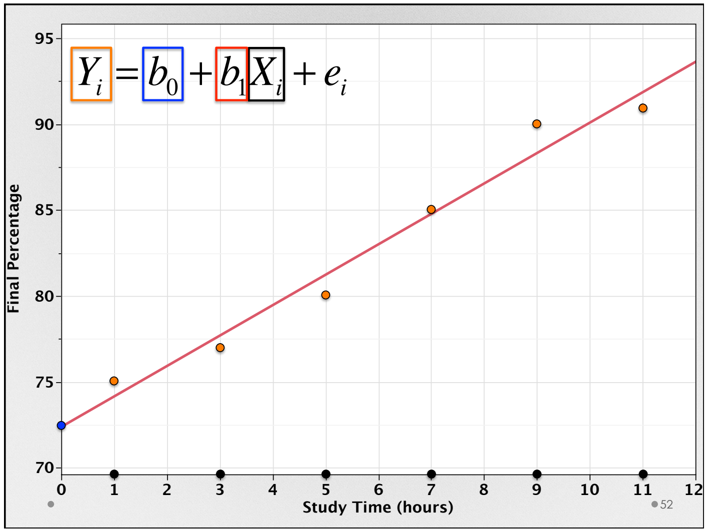

$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
b_1
b_1
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
47

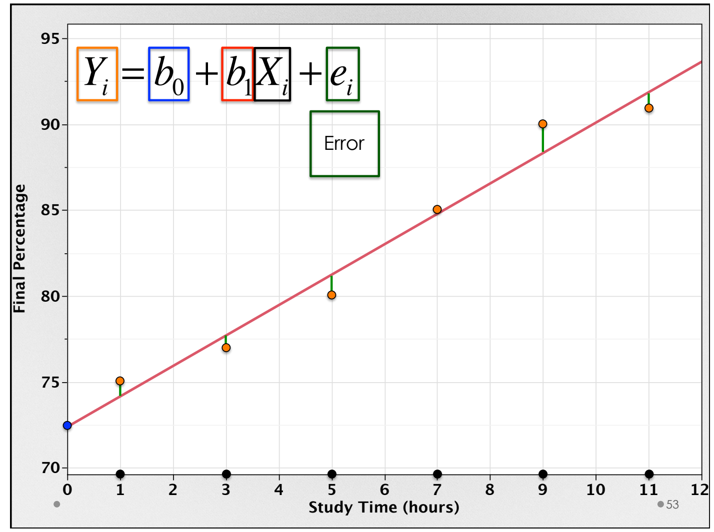

$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
b_1
b_1
Final Percentage
70 75 80 85 90 95
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
48


$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
b_1
b_1
Final Percentage
70 75 80 85 90 95
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
49


$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
b_1
b_1
Final Percentage
70 75 80 85 90 95
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
50


$Y_i = b_0 + b_1 X_i + e_i$
Slope
(Effect)
1
b_1
b_1
Final Percentage
70 75 80 85 90 95
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)
51





One Predictor Linear Regression Model
(Sample Model)

$$\boxed{Y_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i} + \boxed{e_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

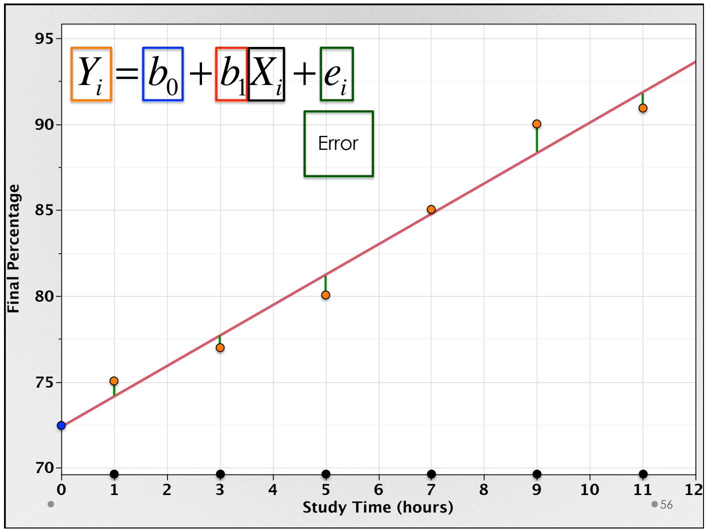
54

"Residual" Error

$$\boxed{e_i} = \boxed{Y_i} - \boxed{\hat{Y}_i}$$

An individual's **residual** is the difference between that individual's **observed value** and the value **predicted by the model**

55



One Predictor Linear Regression Model
(Sample Model)

$$\boxed{Y_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i} + \boxed{e_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

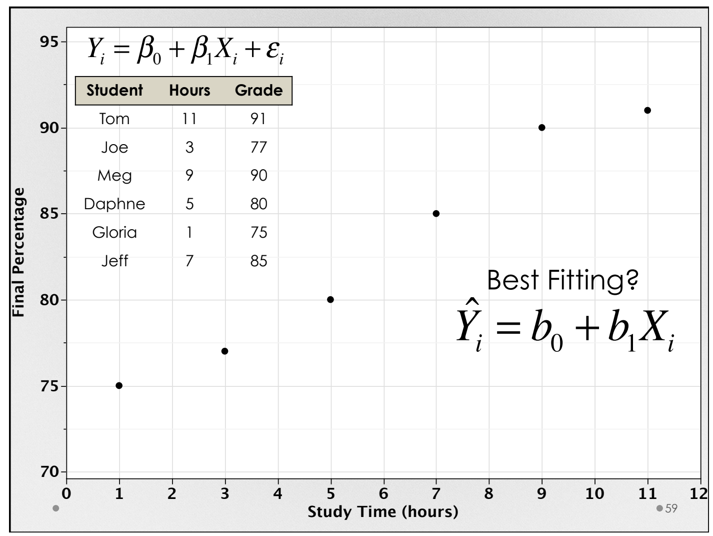
57

One Predictor Linear Regression Model
(Sample Model – Prediction for Y_i)

$$\boxed{\hat{Y}_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i}$$

Predicted Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual)

58



Student	Hours	Grade
Tom	11	91
Joe	3	77
Meg	9	90
Daphne	5	80
Gloria	1	75
Jeff	7	85

Carl Friedrich Gauss

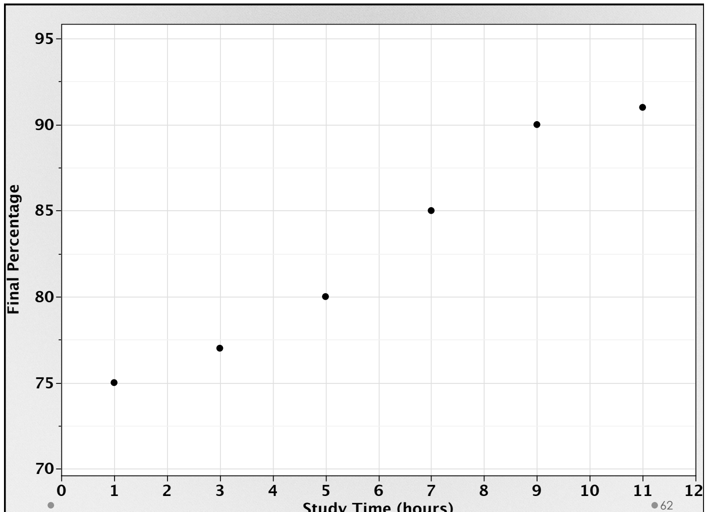
1777 - 1855

Least Squares

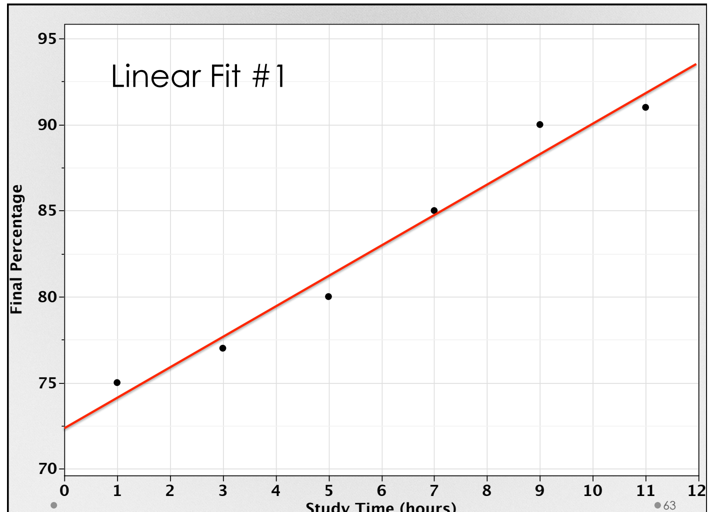
A approach to fitting a model (line) to data where the sum of the squared distances to the data is minimized

☑ Always Produces Unbiased Estimators

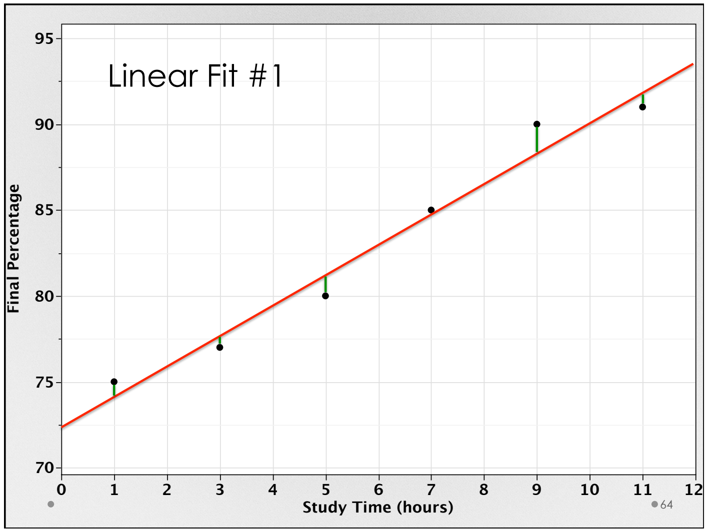
61

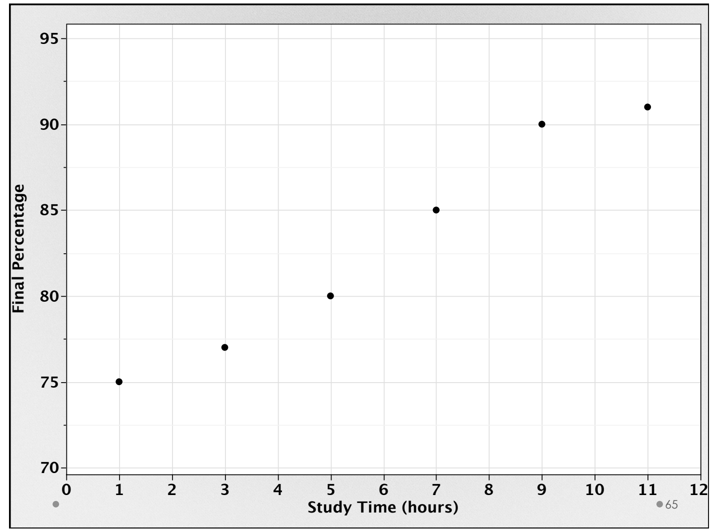


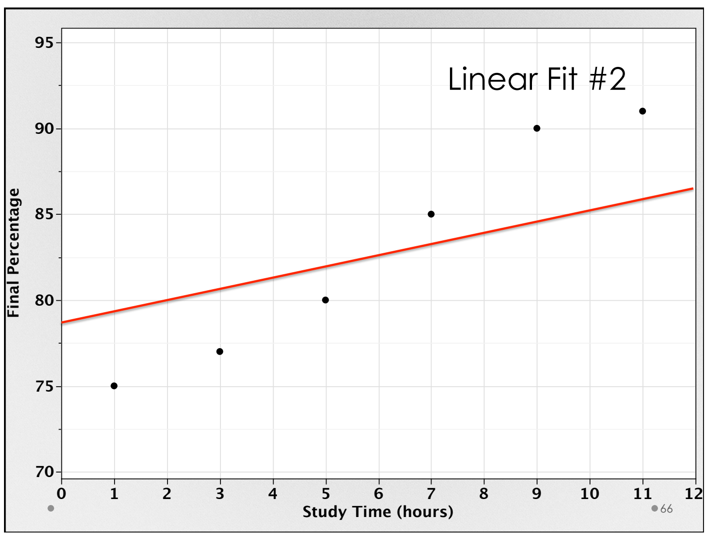
62

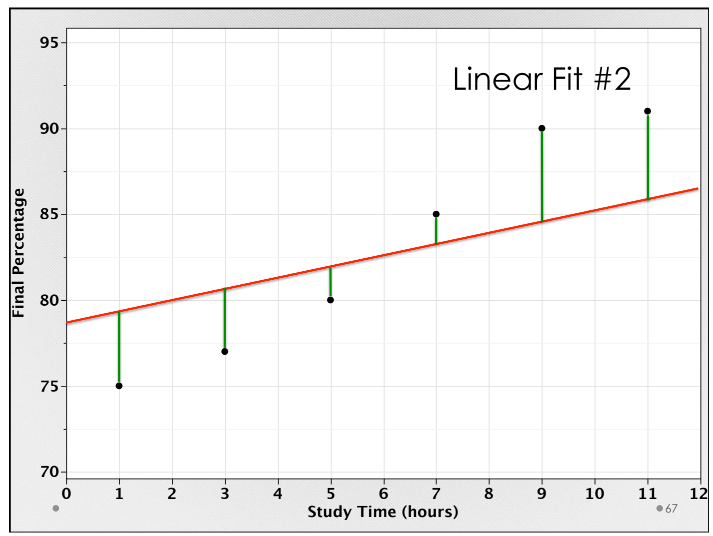


63

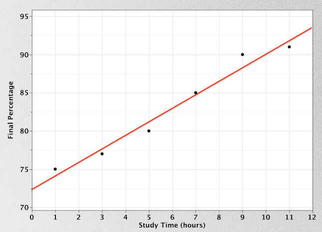

Linear Fit #1
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)


Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)

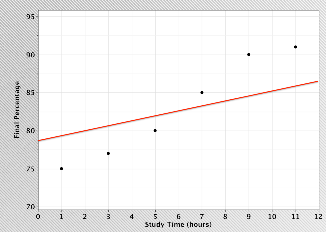

Linear Fit #2
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)


Linear Fit #2
Final Percentage
95
90
85
80
75
70
0 1 2 3 4 5 6 7 8 9 10 11 12
Study Time (hours)

Linear Fit #1



Linear Fit #2



68

Least Squares

A approach to fitting a model (line) to data where the sum of the squared distances to the data is minimized

☑ Always Produces Unbiased Estimators

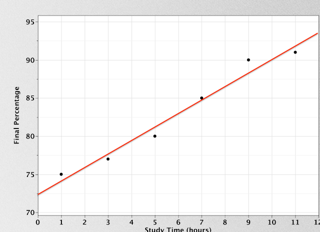
69

Linear Least Squares Criterion

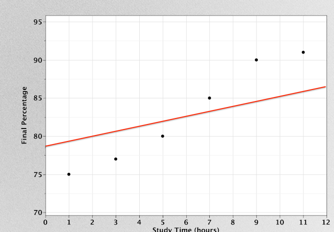
$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

70

Linear Fit #1

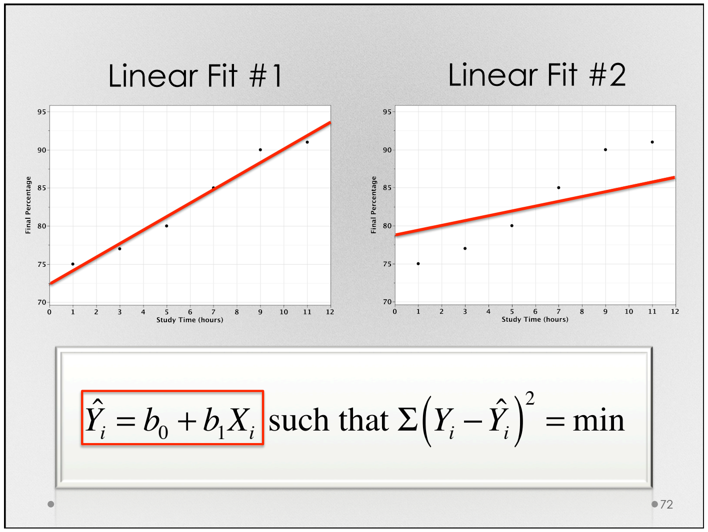


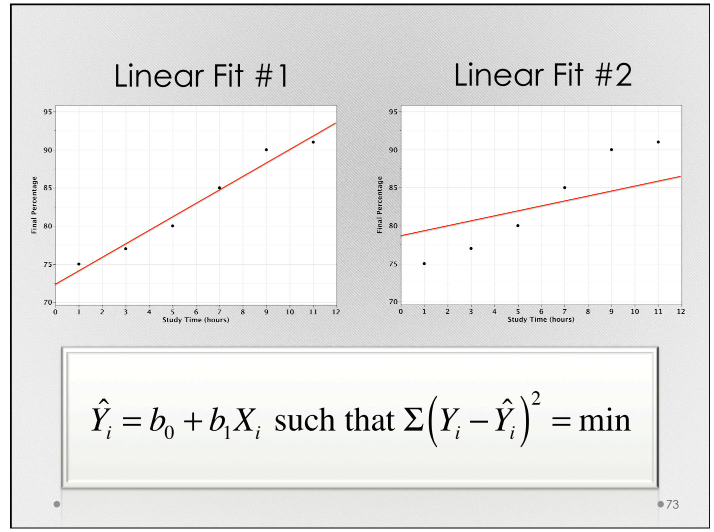
Linear Fit #2

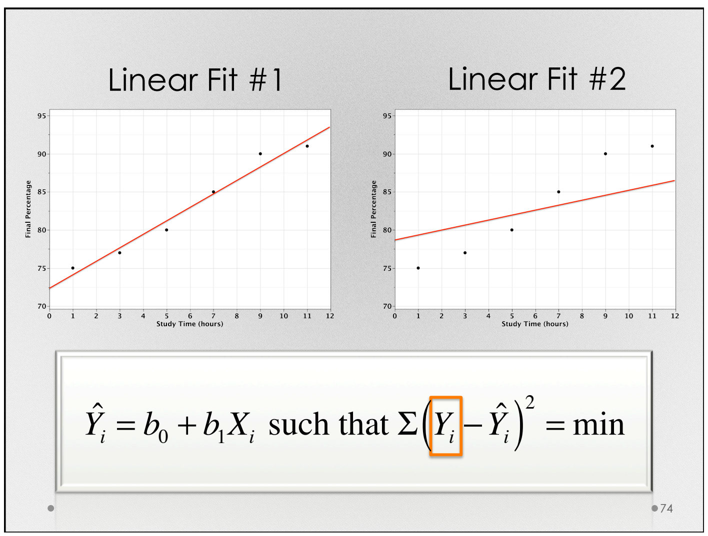


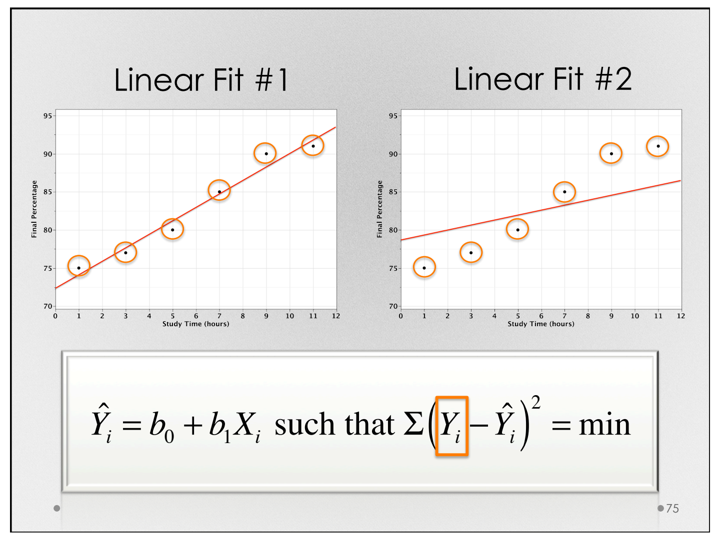
$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

71

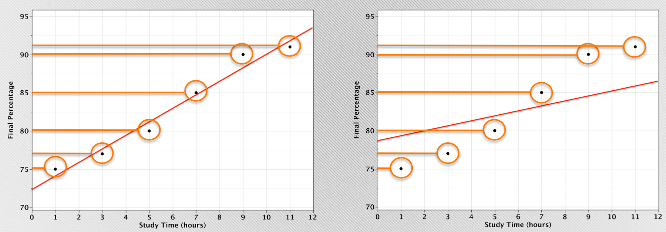

Linear Fit #1
Linear Fit #2
$\boxed{\hat{Y}_i = b_0 + b_1 X_i}$ such that $\Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$
72


Linear Fit #1
Linear Fit #2
$\hat{Y}_i = b_0 + b_1 X_i$ such that $\Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$
73


Linear Fit #1
Linear Fit #2
$\hat{Y}_i = b_0 + b_1 X_i$ such that $\Sigma\left(\boxed{Y_i} - \hat{Y}_i\right)^2 = \min$
74


Linear Fit #1
Linear Fit #2
$\hat{Y}_i = b_0 + b_1 X_i$ such that $\Sigma\left(\boxed{Y_i} - \hat{Y}_i\right)^2 = \min$
75

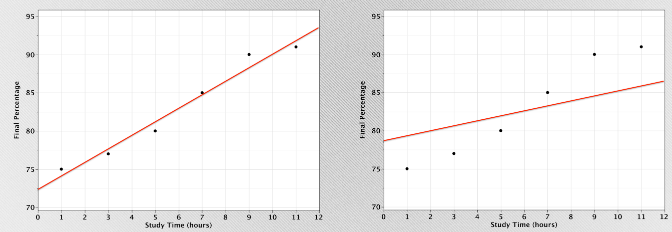
Linear Fit #1 | Linear Fit #2



$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(\boxed{Y_i} - \hat{Y}_i\right)^2 = \min$$

76

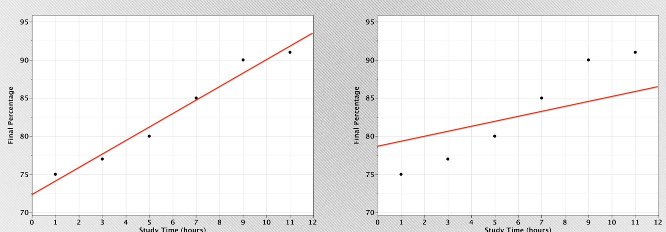
Linear Fit #1 | Linear Fit #2



$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

77

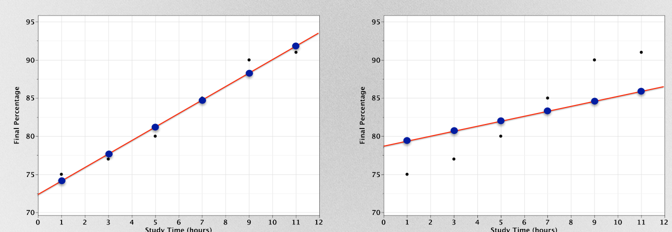
Linear Fit #1 | Linear Fit #2



$$\boxed{\hat{Y}_i} = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \boxed{\hat{Y}_i}\right)^2 = \min$$

78

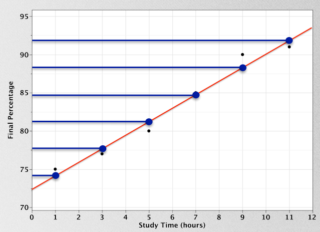
Linear Fit #1 | Linear Fit #2



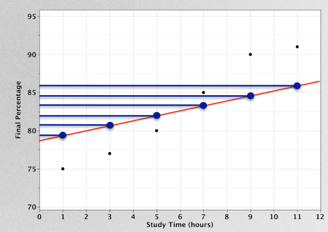
$$\boxed{\hat{Y}_i} = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \boxed{\hat{Y}_i}\right)^2 = \min$$

79

Linear Fit #1



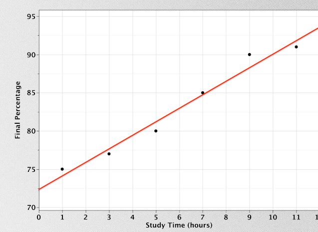
Linear Fit #2



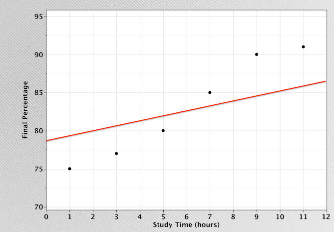
$$\boxed{\hat{Y}_i} = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \boxed{\hat{Y}_i}\right)^2 = \min$$

80

Linear Fit #1



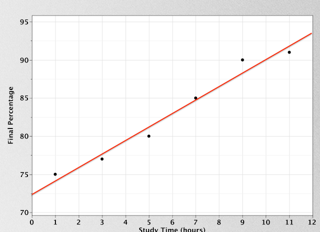
Linear Fit #2



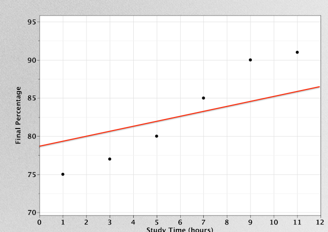
$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

81

Linear Fit #1



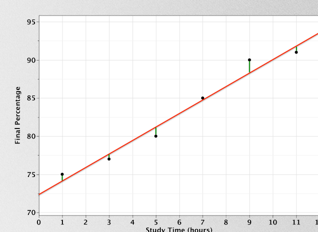
Linear Fit #2



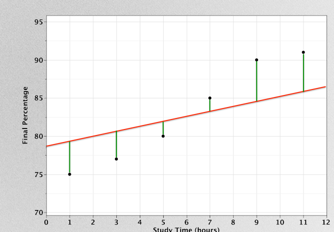
$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\boxed{\left(Y_i - \hat{Y}_i\right)}^2 = \min$$

82

Linear Fit #1



Linear Fit #2



$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\boxed{\left(Y_i - \hat{Y}_i\right)}^2 = \min$$

83

Linear Least Squares Criterion

$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

☑ Always Produces Unbiased Estimators

☑ Closed-Form Solutions

84

Linear Least Squares Criterion

$$\hat{Y}_i = b_0 + b_1 X_i \text{ such that } \Sigma\left(Y_i - \hat{Y}_i\right)^2 = \min$$

☑ Always Produces Unbiased Estimators

☑ Closed-Form Solutions

183

FROM BEFORE

One Predictor Linear Regression Model

(Sample Model – Prediction for Y_i)

$$\boxed{\hat{Y}_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i}$$

Predicted Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual)

184

One Predictor Linear Regression Model

(Population Model – Prediction for Y_i)

$$\boxed{} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i}$$

= Y Intercept + (Slope (Effect) × Score on X for the ith individual)

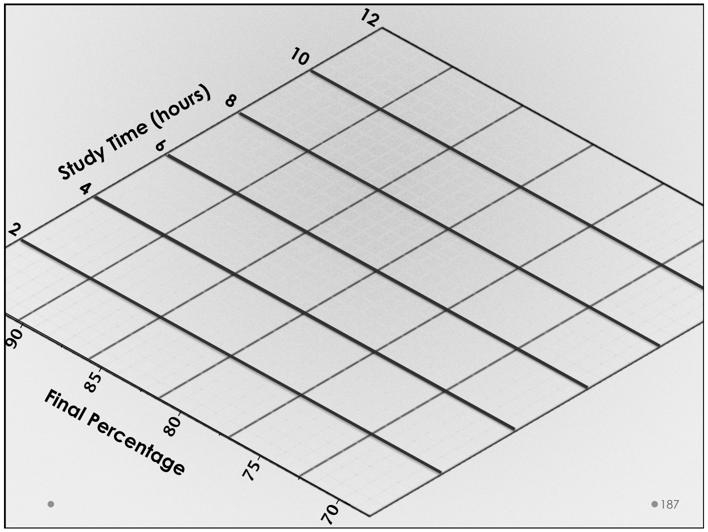
185

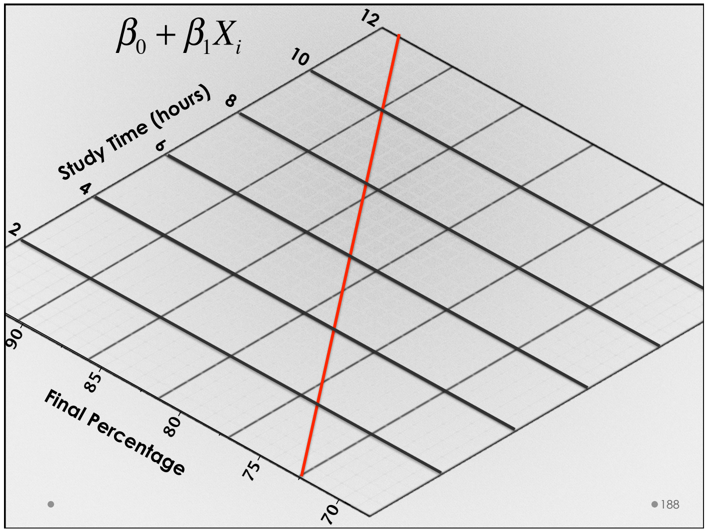
One Predictor Linear Regression Model
(Population Model – Conditional Mean for Y given X_i)

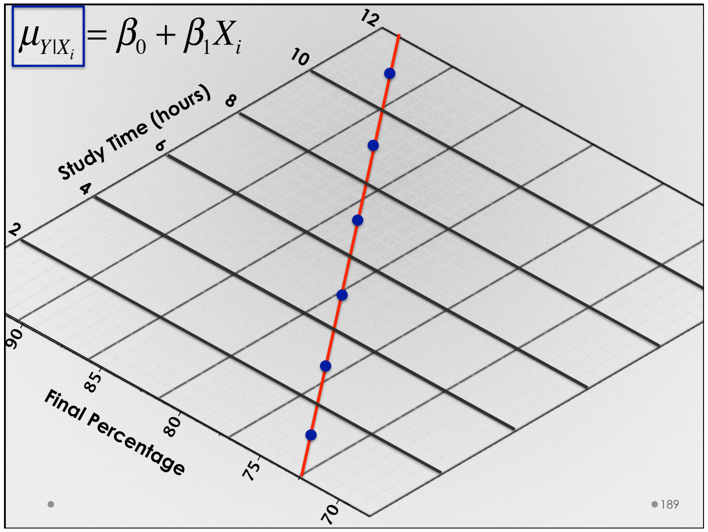
$$\boxed{\mu_{Y|X_i}} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i}$$

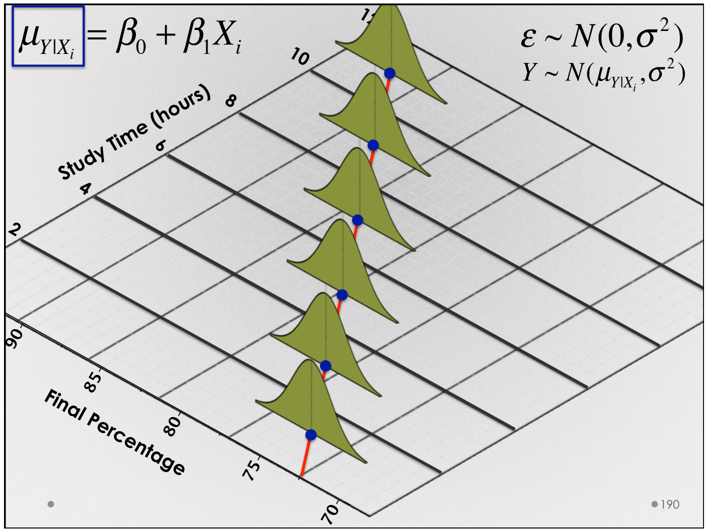
Mean of Y for all Individuals with a particular level of X = Y Intercept + (Slope (Effect) × Score on X for the ith individual)

186









One Predictor Linear Regression Model
(Population Model – Conditional Mean for Y given X_i)

$$\boxed{\mu_{Y|X_i}} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i}$$

Mean of Y for all Individuals with a particular level of X = Y Intercept + (Slope *(Effect)* × Score on X for the ith individual)

191

One Predictor Linear Regression Model
(Sample Model – Prediction for Y_i)

FROM BEFORE

$$\boxed{\hat{Y}_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i}$$

Predicted Score on Y for the ith individual = Y Intercept + (Slope *(Effect)* × Score on X for the ith individual)

192

One Predictor Linear Regression Model
(Population Model – Conditional Mean for Y given X_i)

$$\boxed{\mu_{Y|X_i}} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i}$$

Mean of Y for all Individuals with a particular level of X = Y Intercept + (Slope *(Effect)* × Score on X for the ith individual)

193

One Predictor Linear Regression Model
(Population Model)

$$Y_i = \mu_{Y|X_i} + \varepsilon_i$$

Score on Y for the ith individual = Mean of Y for all Individuals with a particular level of X + Error

194

One Predictor Linear Regression Model
(Population Model)

$$Y_i = \mu_{Y|X_i} + \varepsilon_i$$

Score on Y for the ith individual = Mean of Y for all Individuals with a particular level of X + Error

195

One Predictor Linear Regression Model
(Population Model)

$$Y_i = \underbrace{(\beta_0 + \beta_1 X_i)}_{\mu_{Y|X_i}} + \varepsilon_i$$

Score on Y for the ith individual = (Y Intercept + (Slope *(Effect)* × Score on X for the ith individual)) + Error

196

FROM BEFORE

One Predictor Linear Regression Model
(Population Model)

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Score on Y for the ith individual = Y Intercept + (Slope *(Effect)* × Score on X for the ith individual) + Error

197

One Predictor Linear Regression Model
(Population Model)

FROM BEFORE

$$\boxed{Y_i} = \boxed{\beta_0} + \boxed{\beta_1}\boxed{X_i} + \boxed{\varepsilon_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

198

One Predictor Linear Regression Model
(Sample Model)

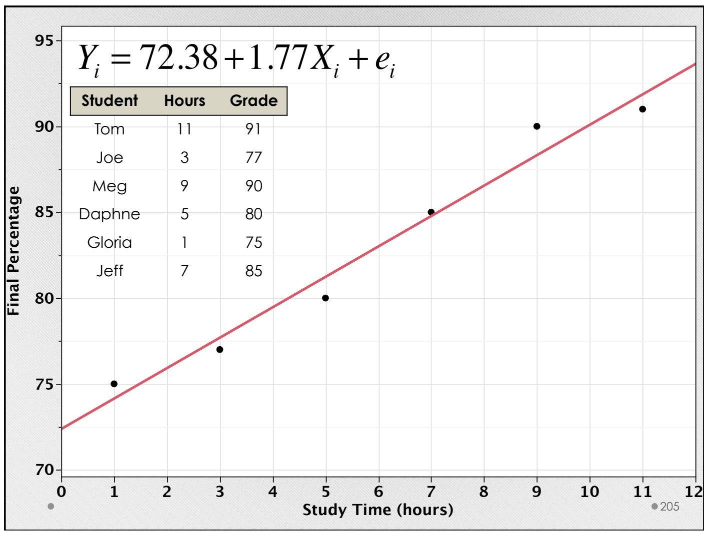
$$\boxed{Y_i} = \boxed{b_0} + \boxed{b_1}\boxed{X_i} + \boxed{e_i}$$

Score on Y for the ith individual = Y Intercept + (Slope (Effect) × Score on X for the ith individual) + Error

204

$$Y_i = 72.38 + 1.77X_i + e_i$$

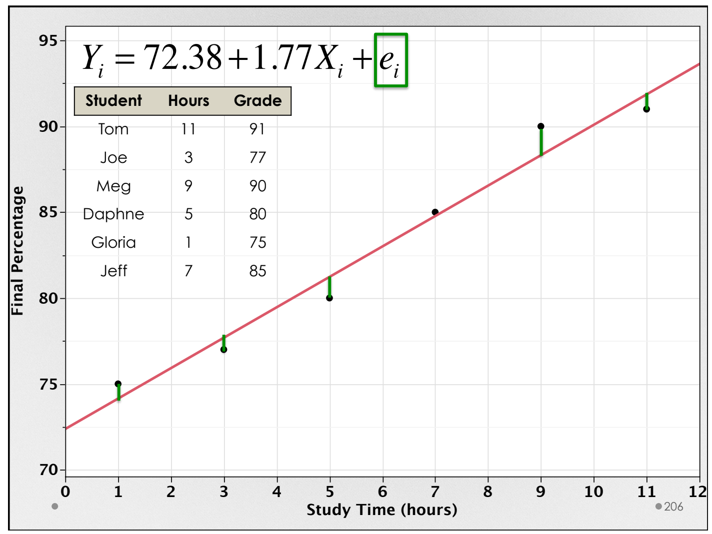
Student	Hours	Grade
Tom	11	91
Joe	3	77
Meg	9	90
Daphne	5	80
Gloria	1	75
Jeff	7	85



205

$$Y_i = 72.38 + 1.77X_i + \boxed{e_i}$$

Student	Hours	Grade
Tom	11	91
Joe	3	77
Meg	9	90
Daphne	5	80
Gloria	1	75
Jeff	7	85



206