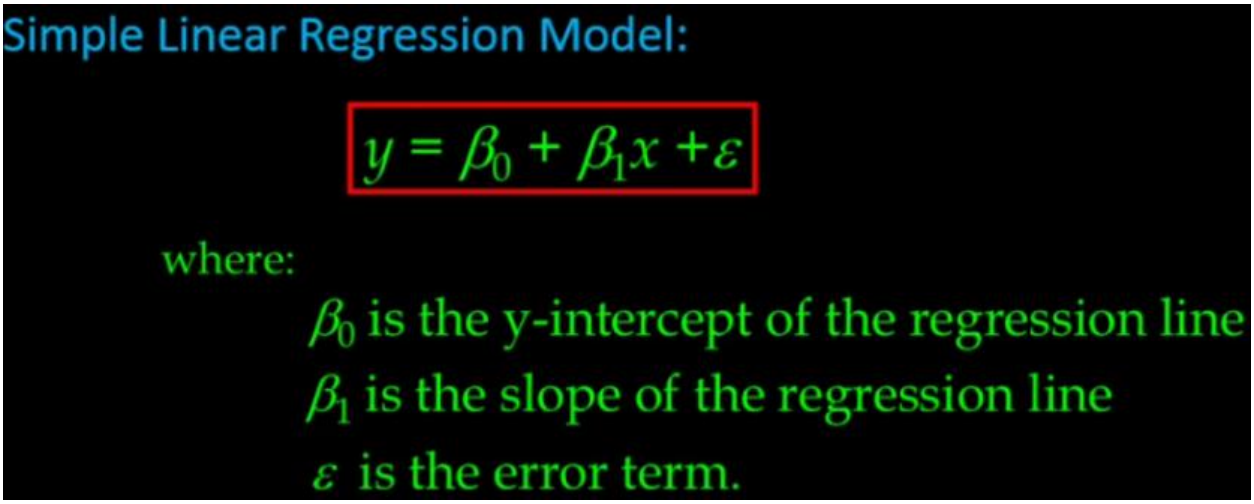
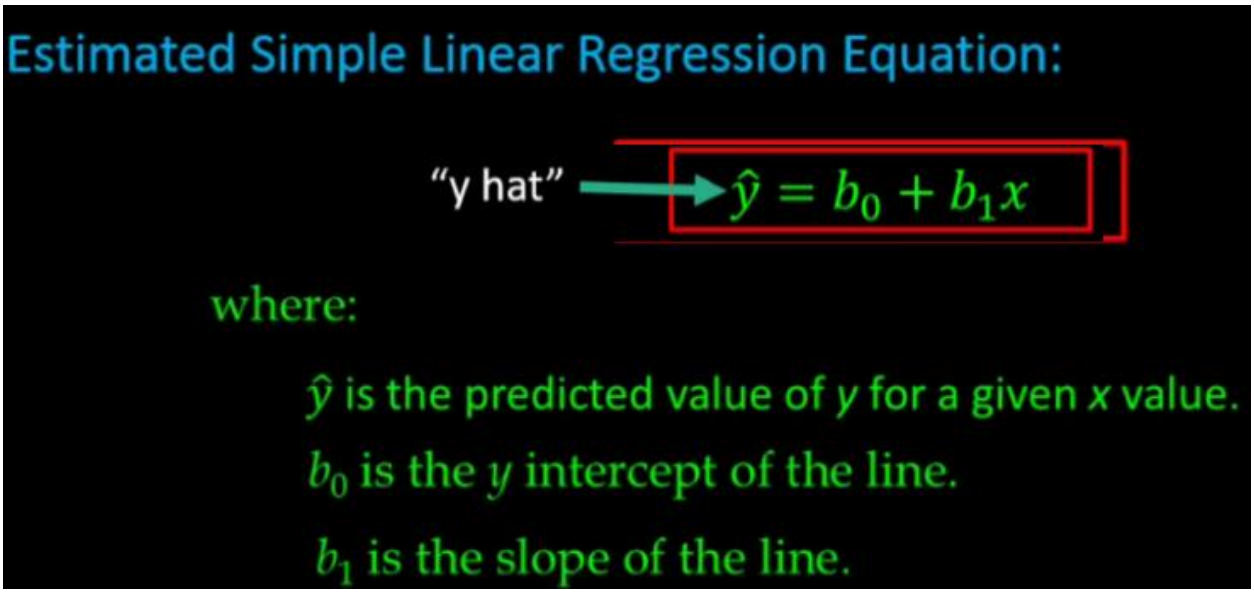
STEPS & FORMULAE FOR REGRESSION ANALYSIS





Least Squares Method:

Use Sample data to find the line of regression

$$\hat{y} = b_0 + b_1 x$$

where:

$\hat{y}$ is the predicted grade on exam

b_0 is the y intercept of the line.

b_1 is the slope of the line.

x is number of hours studied

Least Squares Method:

$$\min \sum (y_i - \hat{y}_i)^2$$

where:

y_i = observed value of the dependent variable for the ith observation

$\hat{y}_i$ = predicted value of the dependent variable for the ith observation

Example: x=3 hours studied

$\hat{y}_i$ = approx. 69

Calculating the Slope:

$$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

where:

x_i = value of independent variable for ith observation

y_i = value of dependent variable for ith observation

$\bar{x}$ = mean value for independent variable

$\bar{y}$ = mean value for dependent variable

Calculating the y – intercept:

$$b_0 = \bar{y} - b_1\bar{x}$$

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
2	69	-2.8	-8.8	24.64	7.84

$\Sigma x_i = 48$	$\Sigma y_i = 778$	320.6	67.6
$\bar{x} = 48/10 = 4.8$	$\bar{y} = 778/10 = 77.8$	$\Sigma(x_i - \bar{x})(y_i - \bar{y})$	$\Sigma(x_i - \bar{x})^2$

$$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$\hat{y} = b_0 + b_1 x$$

Coefficient of Determination:

How well does the regression line fit the data?

$$r^2 = \text{SSR}/\text{SST}$$

where:

SSR = sum of squares due to regression = $\sum(\hat{y}_i - \bar{y})^2$

SST = total sum of squares = $\sum(y_i - \bar{y})^2$

SSE = sum of squares due to error = $\sum(y_i - \hat{y}_i)^2$

SST = SSR + SSE

x_i	y_i	Predicted Grades $\hat{y}_i = 55.048 + 4.74x_i$	Error $y_i - \hat{y}_i$	Squared Error $(y_i - \hat{y}_i)^2$	Deviation $y_i - \bar{y}$	Squared Deviation $(y_i - \bar{y})^2$
2	69	64.528	4.472	19.9988	-8.8	77.44

SSE =79.1215

SST = 1599.6

Coefficient of Determination:

$r^2 = SSR/SST$ $= 1520.4785/1599.6$ $= .9505$

SST = SSR + SSE

SSR = SST - SSE

SSR = 1599.6 - 79.1215

= 1520.4785

r^2 = percent of variability in y can be explained by x

r^2 = 95.05% of the variability in grades can be explained by the number of hours studied