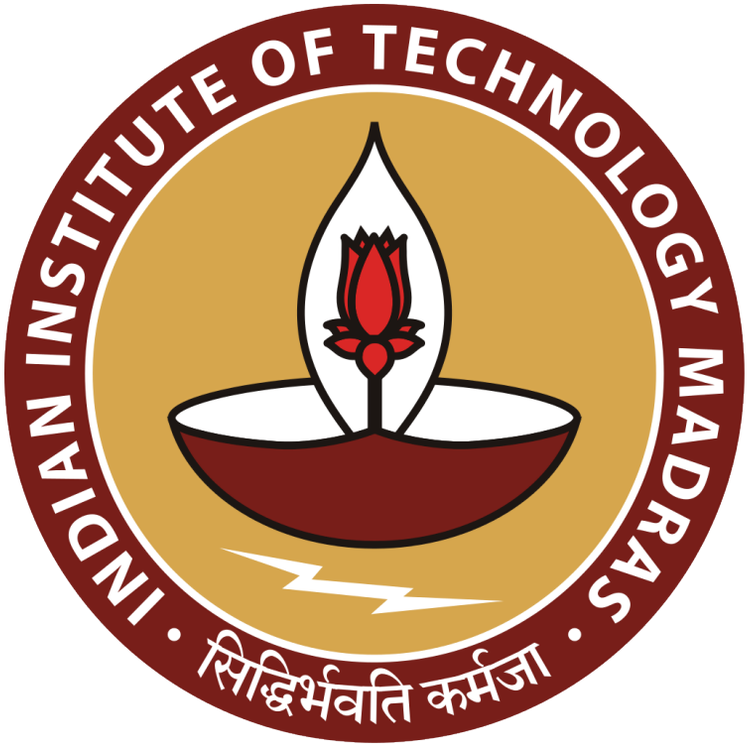


IIT Madras

ONLINE DEGREE

Statistics for Data Science -1

Lecture 4.8: Association between two numerical variables-Fitting a line

Usha Mohan

Indian Institute of Technology Madras

Learning objectives

1. Summarize the linear association between two variables using the equation of a line.
2. Understand the significance of R^2

Summarizing the association with a line

Summarizing the association with a line

- The strength of linear association between the variables was measured using the measures of Covariance and Correlation.

Summarizing the association with a line

- The strength of linear association between the variables was measured using the measures of Covariance and Correlation.
- The linear association can be described using the equation of a line.

Equation of line using google sheets

Equation of line using google sheets

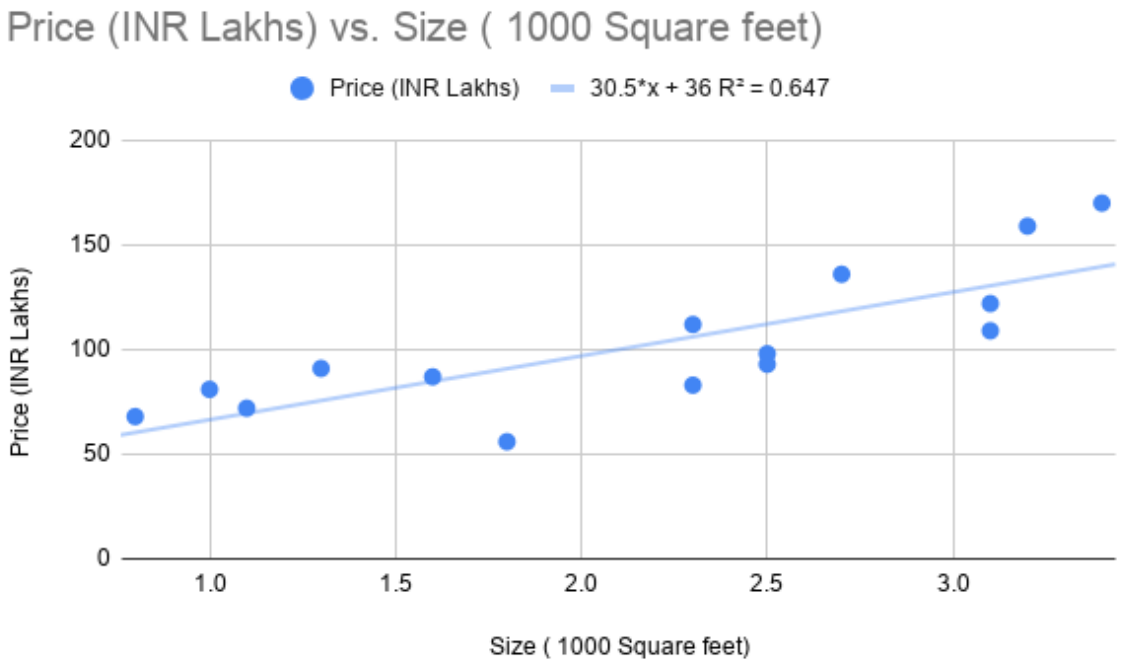
Step 1 Open the scatter plot

Step 2 Under customize tab, click on series

Step 3 Click on trendline

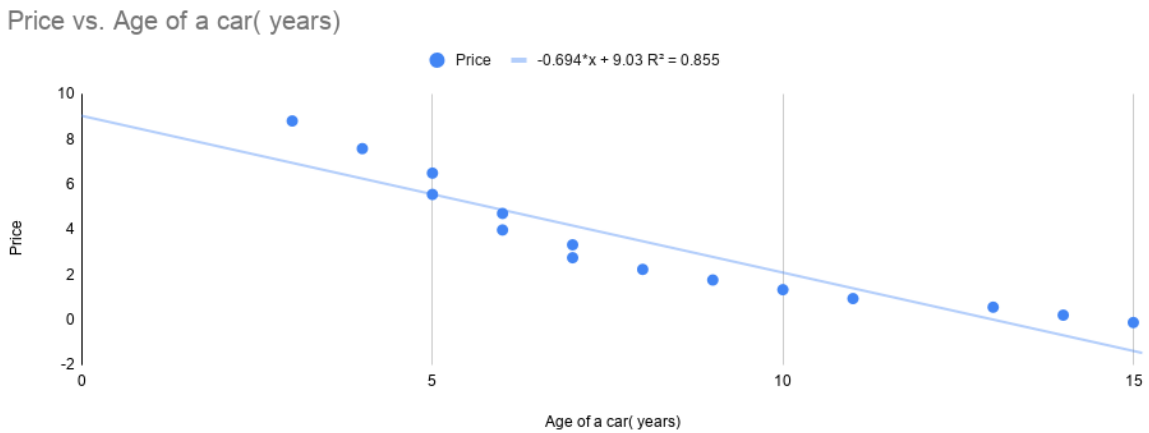
Step 4 Under label tab, click on use equation, and click the show R^2 button.

Example 1: Size versus Price of homes: Equation



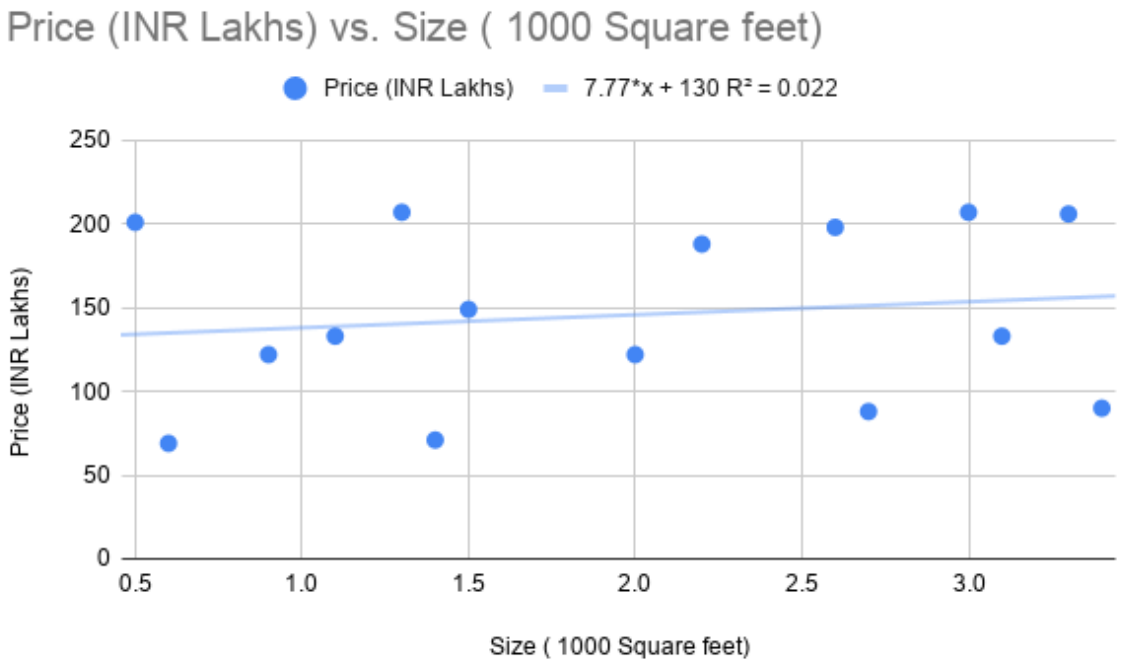
Equation of the line: $Price = 30.5 \times Size + 36$;
$R^2 = 0.647$; $r = 0.804$

Example 2: Age versus Price of cars: Equation



Equation of the line: $Price = -0.694 \times Age + 9.03$;
$R^2 = 0.855$; $r = -0.9247$

Example 3: Size versus Price of homes: Equation



Equation of the line: $Price = 7.77 \times Size + 130$;
$R^2 = 0.022$; $r = 0.149$

Section summary

1. Equation of a line describing linear relationship between two variables.
2. Interpreting slope, R^2 of the line.