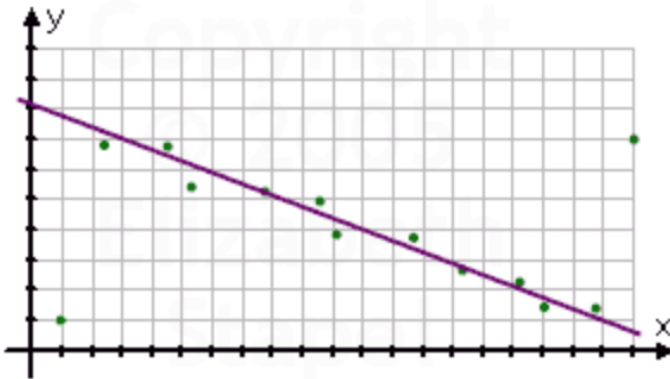


## Session 5: Post Class Test

1. The correlation, covariance and R-squared are all measures that can be used to measure co-movement between two variables.
  - a. What do they share in common?
  - b. What is the difference between correlation and covariance? When might you use one as opposed to the other?
  - c. What is the difference between correlation and R-squared? When might you use one as opposed to the other?
2. When asked to assess whether one variable (X) can explain another (Y), you start with a scatter plot and find a best-fit line through the individual data points.



- a. Using this scatter plot, what can you conclude about the co-movement between these two variables?
  - b. The best-fit line is from an OLS regression. What does OLS stand for and why is it used as the optimizer for fitting the line?
  - c. What are the residual values in a regression? Why do you measure them?
3. In a multiple regression, you still have only one dependent variable, but you can have many independent variables. In picking these independent variables, you would like them to be independent of each other, and each variable should add statistically to the predictive value of your regression.
    - a. How would you check to see if the independent variables are independent of each other? If they are not, what can you do to counter the problem? If you cannot counter the problem, what does it mean for your regression?
    - b. How would you check to see if each variable is adding statistically to your predictive value? If it is not, what would you do?
  4. After you run a regression, you can use it to get predicted values for your dependent variable, given the values of independent variables.
    - a. What are the diagnostic tests that you should run on the predicted values?
    - b. Assuming that you fail these tests, what can you do to fix your problems?
  5. A regression with more observations will have more statistical power and generate better predictions, holding all else constant, than one with fewer observations.
    - a. Explain how sample size plays out in how you structure your regression.
    - b. Explain how sample size plays out in how you read your regression output.