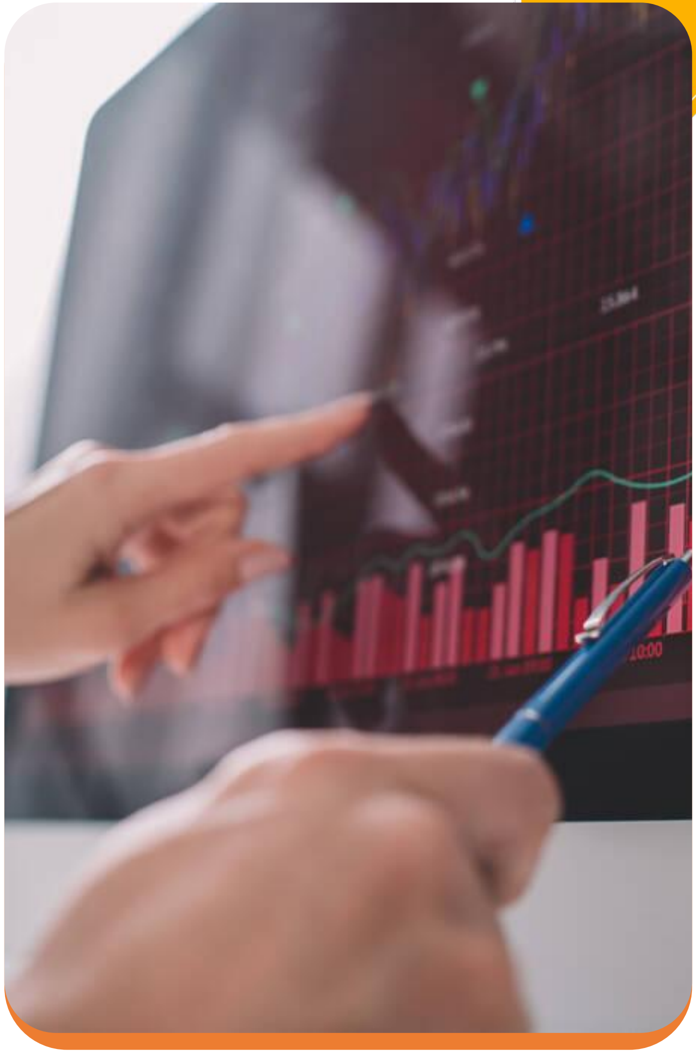
Linear Regression Analysis



Linear Regression Analysis

01 The main purpose of linear regression is to predict the value of a dependent variable based on the values of independent variables.

02 Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables.

Research Questions

- Regression analysis is used to understand the impact of independent variables on a dependent variable.
- Example research question: How does the amount of hours studied affect a student's exam score?

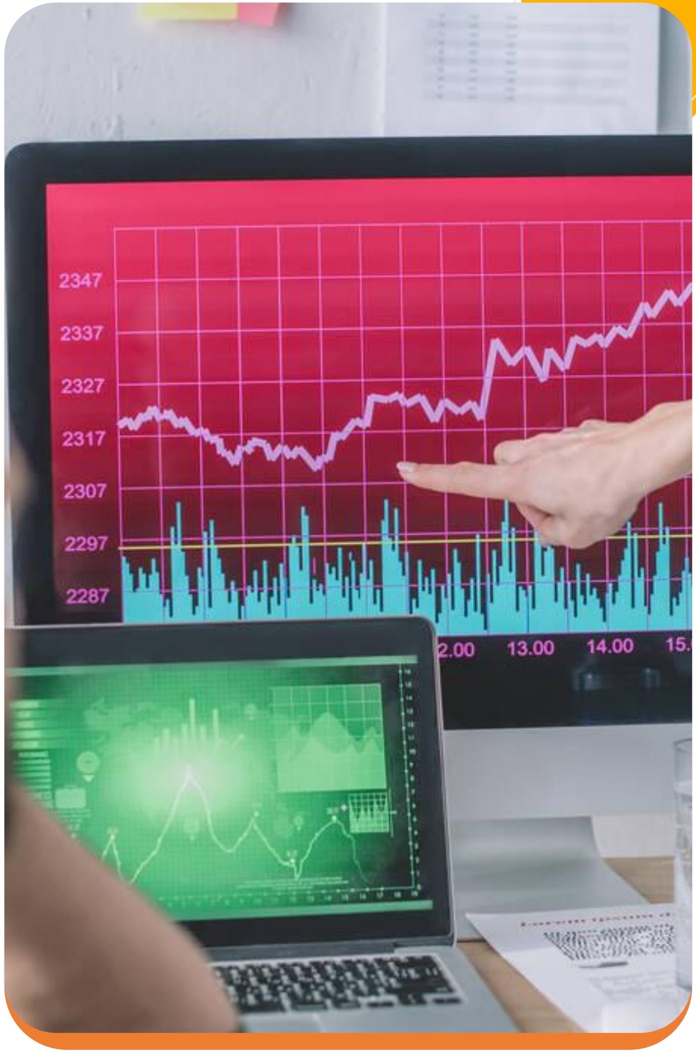
Scenario Example

- Investigating the effect of study hours on exam scores among university students.
- Independent variable: Number of study hours.
- Dependent variable: Exam score.

Independent and Dependent Variables

01 Dependent variable: The outcome variable that the research aims to explain or predict.

02 Independent variables: Variables that influence or predict the outcome.



Simple and Multiple Linear Regression Analysis

01 Multiple Linear Regression: Using two or more independent variables to predict the outcome of a dependent variable.

02 Simple Linear Regression: Modeling the relationship between a single independent variable and a dependent variable.

R-square

- R Squared measures the proportion of the variance in the dependent variable that is predictable from the independent variable(s).
- A higher R Squared value indicates a greater proportion of variance explained by the model.

Beta Value or Regression Coefficient

01 It indicates the strength and direction of the relationship between each independent variable and the dependent variable.

02 Beta value represents the change in the dependent variable for a one-unit change in the independent variable.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon$$

0.5 * 100 = 50%
100%
100% - 50% = 50%

Where:

- Y is the dependent variable (the variable being predicted),
- $X_1, X_2, \ldots, X_n$ are the independent variables (predictors),
- β_0 is the y-intercept of the regression line (it represents the value of Y when all the X variables are 0),
- $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients of the independent variables (they measure the change in the dependent variable Y for a one-unit change in the corresponding independent variable, assuming all other variables are held constant),
- ϵ is the error term (it represents the difference between the observed values and the values predicted by the model; it's a component of the dependent variable that the independent variables do not explain).

Interpretation of the Coefficients:

- **Intercept (β_0)**: This is the expected value of Y when all independent variables are zero. It may not always have a practical interpretation, especially if there's no meaningful condition where all predictors are zero or if such a condition falls outside the scope of the data.
- **Slope Coefficients ($\beta_1, \beta_2, \ldots, \beta_n$)**: Each coefficient represents the expected change in Y for a one-unit change in the corresponding independent variable, holding all other variables constant. For example, if $\beta_1 = 3$, it means that for every one-unit increase in X_1, Y is expected to increase by 3 units, assuming all other variables remain unchanged.

$$\text{Price} = \beta_0 + \beta_1(\overset{10}{\text{Size}}) + \beta_2(\overset{9}{\text{Age}}) + \beta_3(\overset{1}{\text{Garage}}) + \epsilon$$

If $\beta_0 = 50,000$, $\beta_1 = 300$, $\beta_2 = -2,000$, and $\beta_3 = 15,000$, the interpretation would be:

- The base price of a house (with size and age being zero, which is not realistic but serves for interpretation) without a garage is $50,000.
- For each additional square foot of size, the price of the house increases by $300, assuming the age and garage status do not change.
- For each year older the house is, its price decreases by $2,000, assuming size and garage status do not change.
- Having a garage adds $15,000 to the price of the house, assuming size and age remain constant.

Thanks for watching. 😄