# REGRESSION

*Understanding Regression: From Lines to Predictions*

*Presented by Jyothi & Satya*

# What is Regression?

1. Regression is a technique to model the relationship between a **dependent variable** and one or more **independent variables**.
2. Use cases: Forecasting demand, estimating risk, understanding influence.
3. It reveals both **strength** and **direction** of relationships.
4. Enables **predictive modeling** for future outcomes.
5. Types of Regressions :- Simple Regression

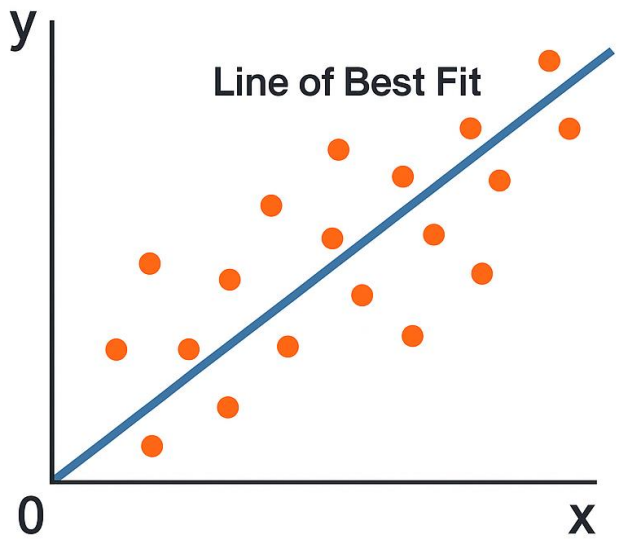
   Multiple Regression

   Logistic Regression

   Polynomial Regression

# Concept of a Line

1. Equation: $y = mx + b \rightarrow$ straight line
2. m: Slope (rate of change)
3. b: Intercept (starting value when $x = 0$)
4. Visual: Line fitted on scatter plot



# Simple Linear Regression

1. Simple linear regression models the relationship between two variables using a straight line.
2. It has one independent variable (x) and one dependent variable (y).
3. The equation is: $\mathbf{y = b_0 + b_1 x}$, where $b_0$ is intercept and $b_1$ is slope.
4. Example: Predicting exam scores based on hours studied.

# Assumptions of Linear Regression

1. Linearity of the relationship
2. Independence of observations
3. Homoscedasticity (constant variance of errors)
4. Normal distribution of residuals

# DIFFERENCE BETWEEN LINEAR GRAPH AND LINE OF BEST FIT

1. A **linear graph** simply refers to any graph where the relationship between variables is represented as a straight line. That could be from a mathematical function like $y = 2x + 3$, or any visually straight plot.
2. A **line of best fit**, on the other hand, is a specific kind of linear graph used in **regression analysis**. It's drawn through a scatter plot of data points to best represent the overall trend—even if the data doesn't perfectly follow a straight line.
3. So: → Every line of best fit on a scatter plot **is a type of linear graph**, → But **not every linear graph is a line of best fit**.

# Multiple Linear Regression

1. **Multiple Linear Regression** models the relationship between one dependent variable and two or more independent variables using a linear equation.
2. Uses **multiple independent variables**
3. Equation: $y = b0 + b1*x1 + b2*x2 + ... + bn*xn$
4. B0 represents the intercept and b1,b2,..bn indicates the coefficients.
5. Example: Predicting house price using size, location, and age

# Applications of Regression

1. Finance: Stock price prediction
2. Healthcare: Predicting disease risk from lifestyle
3. Business: Forecasting sales from marketing budget
4. Education: Predict student performance using attendance, study habits, and test scores.
5. Engineering: Estimate system output or failure risk based on input parameters and operational data.

# Interview Questions

1. What is Regression?
2. What is the difference between simple and multiple linear regression?
3. What are the assumptions of linear regression?
4. What is the role of Intercept in the linear regression?
5. What are the common applications of regression?