

Steps in building regression model

1. Collect/extract data
2. Pre-process it.
3. Creating train and test datasets
4. Visualization and descriptive analytics of patterns present in the data
5. Model building (simple linear regression)
6. Validation and evaluation of model.

The diagram shows the simple linear regression equation $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ with the following labels and arrows:

- Dependent Variable** points to Y_i .
- Population Y intercept** points to β_0 .
- Population Slope Coefficient** points to β_1 .
- Independent Variable** points to X_i .
- Random Error term** points to ϵ_i .

Below the equation, two blue curly braces indicate the components:

- A brace under $\beta_0 + \beta_1 X_i$ is labeled **Linear component**.
- A brace under ϵ_i is labeled **Random Error component**.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x};$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$