Linear Regression (No Code Guide)

1. Introduction to Linear Regression

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

- **Purpose**: It helps in predicting outcomes and understanding how variables are related.

- **Applications**: Business forecasting, risk analysis, economics, medical research, and machine learning.

---

2. Core Concepts

- **Dependent Variable (Y)**: The outcome you want to predict.

- **Independent Variable (X)**: The input or predictor.

- **Line of Best Fit**: A straight line that minimizes the difference between predicted and actual values.

- **Equation**: ($y = mx + c$)

    - (m): slope (effect of X on Y)

    - (c): intercept (value of Y when X = 0)

---

3. Assumptions of Linear Regression

- **Linearity**: Relationship between X and Y is linear.

- **Independence**: Observations are independent of each other.

- **Homoscedasticity**: Constant variance of errors.

- **Normality**: Residuals (errors) are normally distributed.

---

4. Types of Linear Regression

- **Simple Linear Regression**: One independent variable.

- **Multiple Linear Regression**: Two or more independent variables.

---

5. Evaluation Metrics

- **R-squared (R^2)**: Explains how much variance in Y is explained by X.

- **Mean Squared Error (MSE)**: Average squared difference between predicted and actual values.

- **Root Mean Squared Error (RMSE)**: Square root of MSE, easier to interpret.

6. Advantages and Limitations

- **Advantages**:
  - Easy to understand and implement
  - Works well for simple relationships
  - Provides interpretable coefficients

- **Limitations**:
  - Assumes linearity (not suitable for complex patterns)
  - Sensitive to outliers
  - Requires assumptions to be met for reliable results

7. Conclusion

Linear regression is a **foundational tool in statistics and machine learning**. While simple, it provides powerful insights into relationships between variables and serves as a starting point for more advanced modeling techniques.